

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁷ : C07H 21/04, C07K 1/00, 14/00, C12N 1/21, 15/00, 15/09, 15/63, 15/70, C12P 19/34	A1	(11) International Publication Number: WO 00/52027 (43) International Publication Date: 8 September 2000 (08.09.00)
(21) International Application Number: PCT/US00/05432 (22) International Filing Date: 2 March 2000 (02.03.00) (30) Priority Data: 60/122,389 2 March 1999 (02.03.99) US 60/126,049 23 March 1999 (23.03.99) US 60/136,744 28 May 1999 (28.05.99) US (71) Applicant: LIFE TECHNOLOGIES, INC. [US/US]; 9800 Medical Center Drive, Rockville, MD 20850 (US). (72) Inventors: HARTLEY, James, L.; 7409 Hillside Drive, Frederick, MD 21702 (US). BRASCH, Michael, A.; 20931 Sunnycres Road, Gaithersburg, MD 20882 (US). TEMPLE, Gary, F.; 114 Ridge Road, Washington Grove, MD 20882 (US). CHEO, David; 2006 Baltimore Road, #21, Rockville, MD 20851 (US). (74) Agents: ESMOND, Robert, W. et al.; Sterne, Kessler, Goldstein & Fox P.L.L.C., Suite 600, 1100 New York Avenue, N.W., Washington, DC 20005-3934 (US).	(81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG). Published <i>With international search report.</i> <i>With an indication in relation to deposited biological material furnished under Rule 13bis separately from the description.</i>	
(54) Title: COMPOSITIONS AND METHODS FOR USE IN RECOMBINATIONAL CLONING OF NUCLEIC ACIDS (57) Abstract <p>The present invention relates generally to compositions and methods for use in recombinational cloning of nucleic acid molecules. In particular, the invention relates to nucleic acid molecules encoding one or more recombination sites or portions thereof, to nucleic acid molecules comprising one or more of these recombination site nucleotide sequences and optionally comprising one or more additional physical or functional nucleotide sequences. The invention also relates to vectors comprising the nucleic acid molecules of the invention, to host cells comprising the vectors or nucleic acid molecules of the invention, to methods of producing polypeptides using the nucleic acid molecules of the invention, and to polypeptides encoded by these nucleic acid molecules or produced by the methods of the invention. The invention also relates to antibodies that bind to one or more polypeptides of the invention or epitopes thereof. The invention also relates to the use of these compositions in methods for recombinational cloning of nucleic acids, <i>in vitro</i> and <i>in vivo</i>, to provide chimeric DNA molecules that have particular characteristics and/or DNA segments.</p>		

BEST AVAILABLE COPY

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav	TM	Turkmenistan
BF	Burkina Faso	GR	Greece		Republic of Macedonia	TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's	NZ	New Zealand		
CM	Cameroon		Republic of Korea	PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

Compositions and Methods for Use in Recombinational Cloning of Nucleic Acids

BACKGROUND OF THE INVENTION

Field of the Invention

The present invention relates generally to recombinant DNA technology. More particularly, the present invention relates to compositions and methods for use in recombinational cloning of nucleic acid molecules. The invention relates specifically to nucleic acid molecules encoding one or more recombination sites or one or more partial recombination sites, particularly *attB*, *attP*, *attL*, and *attR*, and fragments, mutants, variants and derivatives thereof. The invention also relates to such nucleic acid molecules wherein the one or more recombination site nucleotide sequences is operably linked to the one or more additional physical or functional nucleotide sequences. The invention also relates to vectors comprising the nucleic acid molecules of the invention, to host cells comprising the vectors or nucleic acid molecules of the invention, to methods of producing polypeptides and RNAs encoded by the nucleic acid molecules of the invention, and to polypeptides encoded by these nucleic acid molecules or produced by the methods of the invention, which may be fusion proteins. The invention also relates to antibodies that bind to one or more polypeptides of the invention or epitopes thereof, which may be monoclonal or polyclonal antibodies. The invention also relates to the use of these nucleic acid molecules, vectors, polypeptides and antibodies in methods for recombinational cloning of nucleic acids, *in vitro* and *in vivo*, to provide chimeric DNA molecules that have particular characteristics and/or DNA segments. More particularly, the antibodies of the invention may be used to identify and/or purify proteins or fusion proteins encoded by the nucleic acid molecules or vectors of the invention, or to identify and/or purify the nucleic acid molecules of the invention.

Related Art

Site-specific recombinases. Site-specific recombinases are proteins that are present in many organisms (e.g. viruses and bacteria) and have been characterized to have both endonuclease and ligase properties. These recombinases (along with associated proteins in some cases) recognize specific sequences of bases in DNA and exchange the DNA segments flanking those segments. The recombinases and associated proteins are collectively referred to as "recombination proteins" (see, e.g., Landy, A., *Current Opinion in Biotechnology* 3:699-707 (1993)).

Numerous recombination systems from various organisms have been described. See, e.g., Hoess *et al.*, *Nucleic Acids Research* 14(6):2287 (1986); Abremski *et al.*, *J. Biol. Chem.* 261(1):391 (1986); Campbell, *J. Bacteriol.* 174(23):7495 (1992); Qian *et al.*, *J. Biol. Chem.* 267(11):7794 (1992); Araki *et al.*, *J. Mol. Biol.* 225(1):25 (1992); Maeser and Kahnmann *Mol. Gen. Genet.* 230:170-176 (1991); Esposito *et al.*, *Nucl. Acids Res.* 25(18):3605 (1997).

Many of these belong to the integrase family of recombinases (Argos *et al.* *EMBO J.* 5:433-440 (1986); Voziyanov *et al.*, *Nucl. Acids Res.* 27:930 (1999)). Perhaps the best studied of these are the Integrase/*att* system from bacteriophage λ (Landy, A. *Current Opinions in Genetics and Devel.* 3:699-707 (1993)), the Cre/*loxP* system from bacteriophage P1 (Hoess and Abremski (1990) In *Nucleic Acids and Molecular Biology*, vol. 4. Eds.: Eckstein and Lilley, Berlin-Heidelberg: Springer-Verlag; pp. 90-109), and the FLP/FRT system from the *Saccharomyces cerevisiae* 2 μ circle plasmid (Broach *et al.* *Cell* 29:227-234 (1982)).

Backman (U.S. Patent No. 4,673,640) discloses the *in vivo* use of λ recombinase to recombine a protein producing DNA segment by enzymatic site-specific recombination using wild-type recombination sites *attB* and *attP*.

Hasan and Szybalski (*Gene* 56:145-151 (1987)) discloses the use of λ Int recombinase *in vivo* for intramolecular recombination between wild type *attP* and *attB* sites which flank a promoter. Because the orientations of these sites are

inverted relative to each other, this causes an irreversible flipping of the promoter region relative to the gene of interest.

Palazzolo *et al.* *Gene* 88:25-36 (1990), discloses phage lambda vectors having bacteriophage λ arms that contain restriction sites positioned outside a cloned DNA sequence and between wild-type *loxP* sites. Infection of *E. coli* cells that express the Cre recombinase with these phage vectors results in recombination between the *loxP* sites and the *in vivo* excision of the plasmid replicon, including the cloned cDNA.

Pósfai *et al.* (*Nucl. Acids Res.* 22:2392-2398 (1994)) discloses a method for inserting into genomic DNA partial expression vectors having a selectable marker, flanked by two wild-type FRT recognition sequences. FLP site-specific recombinase as present in the cells is used to integrate the vectors into the genome at predetermined sites. Under conditions where the replicon is functional, this cloned genomic DNA can be amplified.

Bebee *et al.* (U.S. Patent No. 5,434,066) discloses the use of site-specific recombinases such as Cre for DNA containing two *loxP* sites for *in vivo* recombination between the sites.

Boyd (*Nucl. Acids Res.* 21:817-821 (1993)) discloses a method to facilitate the cloning of blunt-ended DNA using conditions that encourage intermolecular ligation to a dephosphorylated vector that contains a wild-type *loxP* site acted upon by a Cre site-specific recombinase present in *E. coli* host cells.

Waterhouse *et al.* (WO 93/19172 and *Nucleic Acids Res.* 21 (9):2265 (1993)) disclose an *in vivo* method where light and heavy chains of a particular antibody were cloned in different phage vectors between *loxP* and *loxP 511* sites and used to transfect new *E. coli* cells. Cre, acting in the host cells on the two parental molecules (one plasmid, one phage), produced four products in equilibrium: two different cointegrates (produced by recombination at either *loxP* or *loxP 511* sites), and two daughter molecules, one of which was the desired product.

Schlake & Bode (*Biochemistry* 33:12746-12751 (1994)) discloses an *in vivo* method to exchange expression cassettes at defined chromosomal locations, each flanked by a wild type and a spacer-mutated FRT recombination site. A

double-reciprocal crossover was mediated in cultured mammalian cells by using this FLP/FRT system for site-specific recombination.

Hartley *et al.* (U.S. Patent No. 5,888,732) disclose compositions and methods for recombinational exchange of nucleic acid segments and molecules, including for use in recombinational cloning of a variety of nucleic acid molecules *in vitro* and *in vivo*, using a combination of wildtype and mutated recombination sites and recombination proteins.

Transposases. The family of enzymes, the transposases, has also been used to transfer genetic information between replicons. Transposons are structurally variable, being described as simple or compound, but typically encode the recombinase gene flanked by DNA sequences organized in inverted orientations. Integration of transposons can be random or highly specific. Representatives such as Tn7, which are highly site-specific, have been applied to the *in vivo* movement of DNA segments between replicons (Lucklow *et al.*, *J. Virol.* 67:4566-4579 (1993)).

Devine and Boeke *Nucl. Acids Res.* 22:3765-3772 (1994), discloses the construction of artificial transposons for the insertion of DNA segments, *in vitro*, into recipient DNA molecules. The system makes use of the integrase of yeast TY1 virus-like particles. The DNA segment of interest is cloned, using standard methods, between the ends of the transposon-like element TY1. In the presence of the TY1 integrase, the resulting element integrates randomly into a second target DNA molecule.

Recombination Sites. Also key to the integration/recombination reactions mediated by the above-noted recombination proteins and/or transposases are recognition sequences, often termed "recombination sites," on the DNA molecules participating in the integration/recombination reactions. These recombination sites are discrete sections or segments of DNA on the participating nucleic acid molecules that are recognized and bound by the recombination proteins during the initial stages of integration or recombination. For example, the recombination site for Cre recombinase is *loxP* which is a 34 base pair sequence comprised of two 13 base pair inverted repeats (serving as the recombinase binding sites) flanking an 8 base pair core sequence. See Figure 1 of Sauer, B., *Curr. Opin. Biotech.*

5:521-527 (1994). Other examples of recognition sequences include the *attB*, *attP*, *attL*, and *attR* sequences which are recognized by the recombination protein λ Int. *attB* is an approximately 25 base pair sequence containing two 9 base pair core-type Int binding sites and a 7 base pair overlap region, while *attP* is an approximately 240 base pair sequence containing core-type Int binding sites and arm-type Int binding sites as well as sites for auxiliary proteins integration host factor (IHF), FIS and excisionase (Xis). See Landy, *Curr. Opin. Biotech.* 3:699-707 (1993); see also U.S. Patent No. 5,888,732, which is incorporated by reference herein.

DNA cloning. The cloning of DNA segments currently occurs as a daily routine in many research labs and as a prerequisite step in many genetic analyses. The purpose of these clonings is various, however, two general purposes can be considered: (1) the initial cloning of DNA from large DNA or RNA segments (chromosomes, YACs, PCR fragments, mRNA, etc.), done in a relative handful of known vectors such as pUC, pGem, pBlueScript, and (2) the subcloning of these DNA segments into specialized vectors for functional analysis. A great deal of time and effort is expended both in the transfer of DNA segments from the initial cloning vectors to the more specialized vectors. This transfer is called subcloning.

The basic methods for cloning have been known for many years and have changed little during that time. A typical cloning protocol is as follows:

- (1) digest the DNA of interest with one or two restriction enzymes;
- (2) gel purify the DNA segment of interest when known;
- (3) prepare the vector by cutting with appropriate restriction enzymes, treating with alkaline phosphatase, gel purify etc., as appropriate;
- (4) ligate the DNA segment to the vector, with appropriate controls to eliminate background of uncut and self-ligated vector;
- (5) introduce the resulting vector into an *E. coli* host cell;
- (6) pick selected colonies and grow small cultures overnight;
- (7) make DNA minipreps; and

(8) analyze the isolated plasmid on agarose gels (often after diagnostic restriction enzyme digestions) or by PCR.

The specialized vectors used for subcloning DNA segments are functionally diverse. These include but are not limited to: vectors for expressing nucleic acid molecules in various organisms; for regulating nucleic acid molecule expression; for providing tags to aid in protein purification or to allow tracking of proteins in cells; for modifying the cloned DNA segment (*e.g.*, generating deletions); for the synthesis of probes (*e.g.*, riboprobes); for the preparation of templates for DNA sequencing; for the identification of protein coding regions; for the fusion of various protein-coding regions; to provide large amounts of the DNA of interest, *etc.* It is common that a particular investigation will involve subcloning the DNA segment of interest into several different specialized vectors.

As known in the art, simple subclonings can be done in one day (*e.g.*, the DNA segment is not large and the restriction sites are compatible with those of the subcloning vector). However, many other subclonings can take several weeks, especially those involving unknown sequences, long fragments, toxic genes, unsuitable placement of restriction sites, high backgrounds, impure enzymes, *etc.* Subcloning DNA fragments is thus often viewed as a chore to be done as few times as possible.

Several methods for facilitating the cloning of DNA segments have been described, *e.g.*, as in the following references.

Ferguson, J., *et al. Gene* 16:191 (1981), discloses a family of vectors for subcloning fragments of yeast DNA. The vectors encode kanamycin resistance. Clones of longer yeast DNA segments can be partially digested and ligated into the subcloning vectors. If the original cloning vector conveys resistance to ampicillin, no purification is necessary prior to transformation, since the selection will be for kanamycin.

Hashimoto-Gotoh, T., *et al. Gene* 41:125 (1986), discloses a subcloning vector with unique cloning sites within a streptomycin sensitivity gene; in a streptomycin-resistant host, only plasmids with inserts or deletions in the dominant sensitivity gene will survive streptomycin selection.

Accordingly, traditional subcloning methods, using restriction enzymes and ligase, are time consuming and relatively unreliable. Considerable labor is expended, and if two or more days later the desired subclone can not be found among the candidate plasmids, the entire process must then be repeated with alternative conditions attempted. Although site specific recombinases have been used to recombine DNA *in vivo*, the successful use of such enzymes *in vitro* was expected to suffer from several problems. For example, the site specificities and efficiencies were expected to differ *in vitro*; topologically linked products were expected; and the topology of the DNA substrates and recombination proteins was expected to differ significantly *in vitro* (see, e.g., Adams *et al*, *J. Mol. Biol.* 226:661-73 (1992)). Reactions that could go on for many hours *in vivo* were expected to occur in significantly less time *in vitro* before the enzymes became inactive. In addition, the stabilities of the recombination enzymes after incubation for extended periods of time in *in vitro* reactions was unknown, as were the effects of the topologies (*i.e.*, linear, coiled, supercoiled, etc.) of the nucleic acid molecules involved in the reaction. Multiple DNA recombination products were expected in the biological host used, resulting in unsatisfactory reliability, specificity or efficiency of subcloning. Thus, *in vitro* recombination reactions were not expected to be sufficiently efficient to yield the desired levels of product.

Accordingly, there is a long felt need to provide an alternative subcloning system that provides advantages over the known use of restriction enzymes and ligases.

SUMMARY OF THE INVENTION

The present invention relates to nucleic acid molecules encoding one or more recombination sites or one or more partial recombination sites, particularly *attB*, *attP*, *attL*, and *attR*, and fragments, mutants, variants and derivatives thereof. The invention also relates to such nucleic acid molecules comprising one or more of the recombination site nucleotide sequences or portions thereof and one or more additional physical or functional nucleotide sequences, such as those

encoding one or more multiple cloning sites, one or more transcription termination sites, one or more transcriptional regulatory sequences (*e.g.*, one or more promoters, enhancers, or repressors), one or more translational signal sequences, one or more nucleotide sequences encoding a fusion partner protein or peptide
5 (*e.g.*, GST, His₆ or thioredoxin), one or more selection markers or modules, one or more nucleotide sequences encoding localization signals such as nuclear localization signals or secretion signals, one or more origins of replication, one or more protease cleavage sites, one or more desired proteins or peptides encoded by a gene or a portion of a gene, and one or more 5' or 3' polynucleotide tails
10 (particularly a poly-G tail). The invention also relates to such nucleic acid molecules wherein the one or more recombination site nucleotide sequences is operably linked to the one or more additional physical or functional nucleotide sequences.

The invention also relates to primer nucleic acid molecules comprising the
15 recombination site nucleotide sequences of the invention (or portions thereof), and to such primer nucleic acid molecules linked to one or more target-specific (*e.g.*, one or more gene-specific) primer nucleic acid sequences. Such primers may also comprise sequences complementary or homologous to DNA or RNA sequences to be amplified, *e.g.*, by PCR, RT-PCR, etc. Such primers may also comprise
20 sequences or portions of sequences useful in the expression of protein genes (ribosome binding sites, localization signals, protease cleavage sites, repressor binding sites, promoters, transcription stops, stop codons, etc.). Said primers may also comprise sequences or portions of sequences useful in the manipulation of DNA molecules (restriction sites, transposition sites, sequencing primers, etc.).
25 The primers of the invention may be used in nucleic acid synthesis and preferably are used for amplification (*e.g.*, PCR) of nucleic acid molecules. When the primers of the invention include target- or gene-specific sequences (any sequence contained within the target to be synthesized or amplified including translation signals, gene sequences, stop codons, transcriptional signals (*e.g.*, promoters) and
30 the like), amplification or synthesis of target sequences or genes may be accomplished. Thus, the invention relates to synthesis of a nucleic acid molecules comprising mixing one or more primers of the invention with a nucleic acid

template, and incubating said mixture under conditions sufficient to make a first nucleic acid molecule complementary to all or a portion of said template. Thus, the invention relates specifically to a method of synthesizing a nucleic acid molecule comprising:

- 5 (a) mixing a nucleic acid template with a polypeptide having polymerase activity and one or more primers comprising one or more recombination sites or portions thereof; and
- (b) incubating said mixture under conditions sufficient to synthesize a first nucleic acid molecule complementary to all or a portion of said
- 10 template and which preferably comprises one or more recombination sites or portions thereof.

Such method of the invention may further comprise incubating said first synthesized nucleic acid molecule under conditions sufficient to synthesize a second nucleic acid molecule complementary to all or a portion of said first nucleic acid molecule. Such synthesis may provide for a first nucleic acid molecule having

15 a recombination site or portion thereof at one or both of its termini.

In a preferred aspect, for the synthesis of the nucleic acid molecules, at least two primers are used wherein each primer comprises a homologous sequence at its terminus and/or within internal sequences of each primer (which may have

20 a homology length of about 2 to about 500 bases, preferably about 3 to about 100 bases, about 4 to about 50 bases, about 5 to about 25 bases and most preferably about 6 to about 18 base overlap). In a preferred aspect, the first such primer comprises at least one target-specific sequence and at least one recombination site or portion thereof while the second primer comprises at least one recombination

25 site or portion thereof. Preferably, the homologous regions between the first and second primers comprise at least a portion of the recombination site. In another aspect, the homologous regions between the first and second primers may comprise one or more additional sequences, *e.g.*, expression signals, translational start motifs, or other sequences adding functionality to the desired nucleic acid

30 sequence upon amplification. In practice, two pairs of primers prime synthesis or amplification of a nucleic acid molecule. In a preferred aspect, all or at least a portion of the synthesized or amplified nucleic acid molecule will be homologous

to all or a portion of the template and further comprises a recombination site or a portion thereof at at least one terminus and preferably both termini of the synthesized or amplified molecule. Such synthesized or amplified nucleic acid molecule may be double stranded or single stranded and may be used in the recombinational cloning methods of the invention. The homologous primers of the invention provide a substantial advantage in that one set of the primers may be standardized for any synthesis or amplification reaction. That is, the primers providing the recombination site sequences (without the target specific sequences) can be pre-made and readily available for use. This in practice allows the use of shorter custom made primers that contain the target specific sequence needed to synthesize or amplify the desired nucleic acid molecule. Thus, this provides reduced time and cost in preparing target specific primers (e.g., shorter primers containing the target specific sequences can be prepared and used in synthesis reactions). The standardized primers, on the other hand, may be produced in mass to reduce cost and can be readily provided (e.g., in kits or as a product) to facilitate synthesis of the desired nucleic acid molecules.

Thus, in one preferred aspect, the invention relates to a method of synthesizing or amplifying one or more nucleic acid molecules comprising:

- (a) mixing one or more nucleic acid templates with at least one polypeptide having polymerase or reverse transcriptase activity and at least a first primer comprising a template specific sequence (complementary to or capable of hybridizing to said templates) and at least a second primer comprising all or a portion of a recombination site wherein said at least a portion of said second primer is homologous to or complementary to at least a portion of said first primer; and
- (b) incubating said mixture under conditions sufficient to synthesize or amplify one or more nucleic acid molecules complementary to all or a portion of said templates and comprising one or more recombination sites or portions thereof at one and preferably both termini of said molecules.

More specifically, the invention relates to a method of synthesizing or amplifying one or more nucleic acid molecules comprising:

- 5 (a) mixing one or more nucleic acid templates with at least one polypeptide having polymerase or reverse transcriptase activity and at least a first primer comprising a template specific sequence (complementary to or capable of hybridizing to said templates) and at least a portion of a recombination site, and at least a second primer comprising all or a portion of a recombination site wherein said at least a portion of said recombination site on said second primer is complementary to or homologous to at least a portion of said recombination site on said first primer; and
- 10 (b) incubating said mixture under conditions sufficient to synthesize or amplify one or more nucleic acid molecules complementary to all or a portion of said templates and comprising one or more recombination sites or portions thereof at one and preferably both termini of said molecules.
- 15

In a more preferred aspect, the invention relates to a method of amplifying or synthesizing one or more nucleic acid molecules comprising:

- 20 (a) mixing one or more nucleic acid templates with at least one polypeptide having polymerase or reverse transcriptase activity and one or more first primers comprising at least a portion of a recombination site and a template specific sequence (complementary to or capable of hybridizing to said template);
- 25 (b) incubating said mixture under conditions sufficient to synthesize or amplify one or more first nucleic acid molecules complementary to all or a portion of said templates wherein said molecules comprise at least a portion of a recombination site at one and preferably both termini of said molecules;
- 30 (c) mixing said molecules with one or more second primers comprising one or more recombination sites, wherein said recombination sites of said second primers are homologous to or

- complementary to at least a portion of said recombination sites on said first nucleic acid molecules; and
- (d) incubating said mixture under conditions sufficient to synthesize or amplify one or more second nucleic acid molecules complementary to all or a portion of said first nucleic acid molecules and which comprise one or more recombination sites at one and preferably both termini of said molecules.

The invention also relates to vectors comprising the nucleic acid molecules of the invention, to host cells comprising the vectors or nucleic acid molecules of the invention, to methods of producing polypeptides encoded by the nucleic acid molecules of the invention, and to polypeptides encoded by these nucleic acid molecules or produced by the methods of the invention, which may be fusion proteins. The invention also relates to antibodies that bind to one or more polypeptides of the invention or epitopes thereof, which may be monoclonal or polyclonal antibodies. The invention also relates to the use of these nucleic acid molecules, primers, vectors, polypeptides and antibodies in methods for recombinational cloning of nucleic acids, *in vitro* and *in vivo*, to provide chimeric DNA molecules that have particular characteristics and/or DNA segments.

The antibodies of the invention may have particular use to identify and/or purify peptides or proteins (including fusion proteins produced by the invention), and to identify and/or purify the nucleic acid molecules of the invention or portions thereof.

The methods for *in vitro* or *in vivo* recombinational cloning of nucleic acid molecule generally relate to recombination between at least a first nucleic acid molecule having at least one recombination site and a second nucleic acid molecule having at least one recombination site to provide a chimeric nucleic acid molecule. In one aspect, the methods relate to recombination between a first vector having at least one recombination site and a second vector having at least one recombination site to provide a chimeric vector. In another aspect, a nucleic acid molecule having at least one recombination site is combined with a vector having at least one recombination site to provide a chimeric vector. In a most preferred aspect, the nucleic acid molecules or vectors used in recombination

comprise two or more recombination sites. In a more specific embodiment of the invention, the recombination methods relate to a Destination Reaction (also referred to herein as an "LR reaction") in which recombination occurs between an Entry clone and a Destination Vector. Such a reaction transfers the nucleic acid molecule of interest from the Entry Clone into the Destination Vector to create an Expression Clone. The methods of the invention also specifically relate to an Entry or Gateward reaction (also referred to herein as a "BP reaction") in which an Expression Clone is recombined with a Donor vector to produce an Entry clone. In other aspects, the invention relates to methods to prepare Entry clones by combining an Entry vector with at least one nucleic acid molecule (e.g., gene or portion of a gene). The invention also relates to conversion of a desired vector into a Destination Vector by including one or more (preferably at least two) recombination sites in the vector of interest. In a more preferred aspect, a nucleic acid molecule (e.g., a cassette) having at least two recombination sites flanking a selectable marker (e.g., a toxic gene or a genetic element preventing the survival of a host cell containing that gene or element, and/or preventing replication, partition or heritability of a nucleic acid molecule (e.g., a vector or plasmid) comprising that gene or element) is added to the vector to make a Destination Vector of the invention.

Preferred vectors for use in the invention include prokaryotic vectors, eukaryotic vectors, or vectors which may shuttle between various prokaryotic and/or eukaryotic systems (e.g. shuttle vectors). Preferred prokaryotic vectors for use in the invention include but are not limited to vectors which may propagate and/or replicate in gram negative and/or gram positive bacteria, including bacteria of the genera *Escherichia*, *Salmonella*, *Proteus*, *Clostridium*, *Klebsiella*, *Bacillus*, *Streptomyces*, and *Pseudomonas* and preferably in the species *E. coli*. Eukaryotic vectors for use in the invention include vectors which propagate and/or replicate in yeast cells, plant cells, mammalian cells, (particularly human and mouse), fungal cells, insect cells, nematode cells, fish cells and the like. Particular vectors of interest include but are not limited to cloning vectors, sequencing vectors, expression vectors, fusion vectors, two-hybrid vectors, gene therapy vectors, phage display vectors, gene-targeting vectors, PACs, BACs, YACs, MACs, and

reverse two-hybrid vectors. Such vectors may be used in prokaryotic and/or eukaryotic systems depending on the particular vector.

In another aspect, the invention relates to kits which may be used in carrying out the methods of the invention, and more specifically relates to cloning
5 or subcloning kits and kits for carrying out the LR Reaction (*e.g.*, making an Expression Clone), for carrying out the BP Reaction (*e.g.*, making an Entry Clone), and for making Entry Clone and Destination Vector molecules of the invention. Such kits may comprise a carrier or receptacle being compartmentalized to receive and hold therein any number of containers. Such containers may
10 contain any number of components for carrying out the methods of the invention or combinations of such components. In particular, a kit of the invention may comprise one or more components (or combinations thereof) selected from the group consisting of one or more recombination proteins or auxiliary factors or combinations thereof, one or more compositions comprising one or more
15 recombination proteins or auxiliary factors or combinations thereof (for example, GATEWAY™ LR Clonase™ Enzyme Mix or GATEWAY™ BP Clonase™ Enzyme Mix), one or more reaction buffers, one or more nucleotides, one or more primers of the invention, one or more restriction enzymes, one or more ligases, one or more polypeptides having polymerase activity (*e.g.*, one or more reverse
20 transcriptases or DNA polymerases), one or more proteinases (*e.g.*, proteinase K or other proteinases), one or more Destination Vector molecules, one or more Entry Clone molecules, one or more host cells (*e.g.* competent cells, such as *E. coli* cells, yeast cells, animal cells (including mammalian cells, insect cells, nematode cells, avian cells, fish cells, etc.), plant cells, and most particularly
25 *E. coli* DB3.1 host cells, such as *E. coli* LIBRARY EFFICIENCY® DB3.1™ Competent Cells), instructions for using the kits of the invention (*e.g.*, to carry out the methods of the invention), and the like. In related aspects, the kits of the invention may comprise one or more nucleic acid molecules encoding one or more recombination sites or portions thereof, particularly one or more nucleic acid
30 molecules comprising a nucleotide sequence encoding the one or more recombination sites or portions thereof of the invention. Preferably, such nucleic acid molecules comprise at least two recombination sites which flank a selectable

marker (*e.g.*, a toxic gene and/or antibiotic resistance gene). In a preferred aspect, such nucleic acid molecules are in the form of a cassette (*e.g.*, a linear nucleic acid molecule comprising one or more and preferably two or more recombination sites or portions thereof).

5 Kits for inserting or adding recombination sites to nucleic acid molecules of interest may comprise one or more nucleases (preferably restriction endonucleases), one or more ligases, one or more topoisomerases, one or more polymerases, and one or more nucleic acid molecules or adapters comprising one or more recombination sites. Kits for integrating recombination sites into one or
10 more nucleic acid molecules of interest may comprise one or more components (or combinations thereof) selected from the group consisting of one or more integration sequences comprising one or more recombination sites. Such integration sequences may comprise one or more transposons, integrating viruses, homologous recombination sequences, RNA molecules, one or more host cells
15 and the like.

 Kits for making the Entry Clone molecules of the invention may comprise any or a number of components and the composition of such kits may vary depending on the specific method involved. Such methods may involve inserting
20 the nucleic acid molecules of interest into an Entry or Donor Vector by the recombinational cloning methods of the invention, or using conventional molecular biology techniques (*e.g.*, restriction enzyme digestion and ligation). In a preferred aspect, the Entry Clone is made using nucleic acid amplification or synthesis products. Kits for synthesizing Entry Clone molecules from amplification or
25 synthesis products may comprise one or more components (or combinations thereof) selected from the group consisting of one or more Donor Vectors (*e.g.*, one or more attP vectors including, but not limited to, pDONR201 (Figure 49), pDONR202 (Figure 50), pDONR203 (Figure 51), pDONR204 (Figure 52), pDONR205 (Figure 53), pDONR206 (Figure 53), and the like), one or more polypeptides having polymerase activity (preferably DNA polymerases and most
30 preferably thermostable DNA polymerases), one or more proteinases, one or more reaction buffers, one or more nucleotides, one or more primers comprising one or

more recombination sites or portions thereof, and instructions for making one or more Entry Clones.

Kits for making the Destination vectors of the invention may comprise any number of components and the compositions of such kits may vary depending on the specific method involved. Such methods may include the recombination methods of the invention or conventional molecular biology techniques (e.g., restriction endonuclease digestion and ligation). In a preferred aspect, the Destination vector is made by inserting a nucleic acid molecule comprising at least one recombination site (or portion thereof) of the invention (preferably a nucleic acid molecule comprising at least two recombination sites or portions thereof flanking a selectable marker) into a desired vector to convert the desired vector into a Destination vector of the invention. Such kits may comprise at least one component (or combinations thereof) selected from the group consisting of one or more restriction endonucleases, one or more ligases, one or more polymerases, one or more nucleotides, reaction buffers, one or more nucleic acid molecules comprising at least one recombination site or portion thereof (preferably at least one nucleic acid molecule comprising at least two recombination sites flanking at least one selectable marker, such as a cassette comprising at least one selectable marker such as antibiotic resistance genes and/or toxic genes), and instructions for making such Destination vectors.

The invention also relates to kits for using the antibodies of the invention in identification and/or isolation of peptides and proteins (which may be fusion proteins) produced by the nucleic acid molecules of the invention, and for identification and/or isolation of the nucleic acid molecules of the invention or portions thereof. Such kits may comprise one or more components (or combination thereof) selected from the group consisting of one or more antibodies of the invention, one or more detectable labels, one or more solid supports and the like.

Other preferred embodiments of the present invention will be apparent to one of ordinary skill in light of what is known in the art, in light of the following drawings and description of the invention, and in light of the claims.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 depicts one general method of the present invention, wherein the starting (parent) DNA molecules can be circular or linear. The goal is to exchange the new subcloning vector D for the original cloning vector B. It is desirable in one embodiment to select for AD and against all the other molecules, including the Cointegrate. The square and circle are sites of recombination: *e.g.*, *lox* (such as *loxP*) sites, *att* sites, *etc.* For example, segment D can contain expression signals, protein fusion domains, new drug markers, new origins of replication, or specialized functions for mapping or sequencing DNA. It should be noted that the cointegrate molecule contains Segment D (Destination vector) adjacent to segment A (Insert), thereby juxtaposing functional elements in D with the insert in A. Such molecules can be used directly in vitro (*e.g.*, if a promoter is positioned adjacent to a gene-for in vitro transcription/translation) or in vivo (following isolation in a cell capable of propagating *ccdB*-containing vectors) by selecting for the selection markers in Segments B+D. As one skilled in the art will recognize, this single step method has utility in certain envisioned applications of the invention.

Figure 2 is a more detailed depiction of the recombinational cloning system of the invention, referred to herein as the "GATEWAY™ Cloning System." This figure depicts the production of Expression Clones via a "Destination Reaction," which may also be referred to herein as an "LR Reaction." A *kan^r* vector (referred to herein as an "Entry clone") containing a DNA molecule of interest (*e.g.*, a gene) localized between an *attL1* site and an *attL2* site is reacted with an *amp^r* vector (referred to herein as a "Destination Vector") containing a toxic or "death" gene localized between an *attR1* site and an *attR2* site, in the presence of GATEWAY™ LR Clonase™ Enzyme Mix (a mixture of Int, IHF and Xis). After incubation at 25°C for about 60 minutes, the reaction yields an *amp^r* Expression Clone containing the DNA molecule of interest localized between an *attB1* site and an *attB2* site, and a *kan^r* byproduct molecule, as well as intermediates. The reaction mixture may then be transformed into host cells (*e.g.*, *E. coli*) and clones containing the nucleic acid molecule of interest may

be selected by plating the cells onto ampicillin-containing media and picking amp^r colonies.

Figure 3 is a schematic depiction of the cloning of a nucleic acid molecule from an Entry clone into multiple types of Destination vectors, to produce a variety of Expression Clones. Recombination between a given Entry clone and different types of Destination vectors (not shown), via the LR Reaction depicted in Figure 2, produces multiple different Expression Clones for use in a variety of applications and host cell types.

Figure 4 is a detailed depiction of the production of Entry Clones via a "BP reaction," also referred to herein as an "Entry Reaction" or a "Gateway Reaction." In the example shown in this figure, an amp^r expression vector containing a DNA molecule of interest (*e.g.*, a gene) localized between an *attB*1 site and an *attB*2 site is reacted with a kan^r Donor vector (*e.g.*, an *attP* vector; here, GATEWAY™ pDONR201 (see Figure 49A-C)) containing a toxic or "death" gene localized between an *attP*1 site and an *attP*2 site, in the presence of GATEWAY™ BP Clonase™ Enzyme Mix (a mixture of Int and IHF). After incubation at 25°C for about 60 minutes, the reaction yields a kan^r Entry clone containing the DNA molecule of interest localized between an *attL*1 site and an *attL*2 site, and an amp^r by-product molecule. The Entry clone may then be transformed into host cells (*e.g.*, *E. coli*) and clones containing the Entry clone (and therefore the nucleic acid molecule of interest) may be selected by plating the cells onto kanamycin-containing media and picking kan^r colonies. Although this figure shows an example of use of a kan^r Donor vector, it is also possible to use Donor vectors containing other selection markers, such as the gentamycin resistance or tetracycline resistance markers, as discussed herein.

Figure 5 is a more detailed schematic depiction of the LR ("Destination") reaction (Figure 5A) and the BP ("Entry" or "Gateway") reaction (Figure 5B) of the GATEWAY™ Cloning System, showing the reactants, products and byproducts of each reaction.

Figure 6 shows the sequences of the attB1 and attB2 sites flanking a gene of interest after subcloning into a Destination Vector to create an Expression Clone.

Figure 7 is a schematic depiction of four ways to make Entry Clones using the compositions and methods of the invention: 1. using restriction enzymes and ligase; 2. starting with a cDNA library prepared in an attL Entry Vector; 3. using an Expression Clone from a library prepared in an attB Expression Vector via the BxP reaction; and 4. recombinational cloning of PCR fragments with terminal attB sites, via the BxP reaction. Approaches 3 and 4 rely on recombination with a Donor vector (here, an attP vector such as pDONR201 (see Figure 49A-C), pDONR202 (see Figure 50A-C), pDONR203 (see Figure 51A-C), pDONR204 (see Figure 52A-C), pDONR205 (see Figure 53A-C), or pDONR206 (see Figure 54A-C), for example) that provides an Entry Clone carrying a selection marker such as kan^r, gen^r, tet^r, or the like.

Figure 8 is a schematic depiction of cloning of a PCR product by a BxP (Entry or Gateward) reaction. A PCR product with 25 bp terminal attB sites (plus four Gs) is shown as a substrate for the BxP reaction. Recombination between the attB-PCR product of a gene and a Donor vector (which donates an Entry Vector that carries kan^r) results in an Entry Clone of the PCR product.

Figure 9 is a listing of the nucleotide sequences of the recombination sites designated herein as attB1, attB2, attP1, attP2, attL1, attL2, attR1 and attR2. Sequences are written conventionally, from 5' to 3'.

Figures 10-20: The plasmid backbone for all the Entry Vectors depicted herein is the same, and is shown in Figure 10A for the Entry Vector pENTR1A. For other Entry Vectors shown in Figures 11-20, only the sequences shown in Figure "A" for each figure set (*i.e.*, Figure 11A, Figure 12A, etc.) are different (within the attL1-attL2 cassettes) from those shown in Figure 10 -- the plasmid backbone is identical.

Figure 10 is a schematic depiction of the physical map and cloning sites (Figure 10A), and the nucleotide sequence (Figure 10B), of the Entry Vector pENTR1A.

Figure 11 is a schematic depiction of the cloning sites (Figure 11A) and the nucleotide sequence (Figure 11B) of the Entry Vector pENTR2B.

Figure 12 is a schematic depiction of the cloning sites (Figure 12A) and the nucleotide sequence (Figure 12B) of the Entry Vector pENTR3C.

5 **Figure 13** is a schematic depiction of the cloning sites (Figure 13A) and the nucleotide sequence (Figure 13B) of the Entry Vector pENTR4.

Figure 14 is a schematic depiction of the cloning sites (Figure 14A) and the nucleotide sequence (Figure 14B) of the Entry Vector pENTR5.

10 **Figure 15** is a schematic depiction of the cloning sites (Figure 15A) and the nucleotide sequence (Figure 15B) of the Entry Vector pENTR6.

Figure 16 is a schematic depiction of the cloning sites (Figure 16A) and the nucleotide sequence (Figure 16B) of the Entry Vector pENTR7.

Figure 17 is a schematic depiction of the cloning sites (Figure 17A) and the nucleotide sequence (Figure 17B) of the Entry Vector pENTR8.

15 **Figure 18** is a schematic depiction of the cloning sites (Figure 18A) and the nucleotide sequence (Figure 18B) of the Entry Vector pENTR9.

Figure 19 is a schematic depiction of the cloning sites (Figure 19A) and the nucleotide sequence (Figure 19B) of the Entry Vector pENTR10.

20 **Figure 20** is a schematic depiction of the cloning sites (Figure 20A) and the nucleotide sequence (Figure 20B) of the Entry Vector pENTR11.

25 **Figure 21** is a schematic depiction of the physical map and the Trc expression cassette (Figure 21A) showing the promoter sequences at -35 and at -10 from the initiation codon, and the nucleotide sequence (Figure 21B-D), of Destination Vector pDEST1. This vector may also be referred to as pTrc-DEST1.

30 **Figure 22** is a schematic depiction of the physical map and the His6 expression cassette (Figure 22A) showing the promoter sequences at -35 and at -10 from the initiation codon, and the nucleotide sequence (Figure 22B-D), of Destination Vector pDEST2. This vector may also be referred to as pHis6-DEST2.

Figure 23 is a schematic depiction of the physical map and the GST expression cassette (Figure 23A) showing the promoter sequences at -35 and at -10 from the initiation codon, and the nucleotide sequence (Figure 23B-D), of Destination Vector pDEST3. This vector may also be referred to as pGST-DEST3.

Figure 24 is a schematic depiction of the physical map and the His6-Trx expression cassette (Figure 24A) showing the promoter sequences at -35 and at -10 from the initiation codon and a TEV protease cleavage site, and the nucleotide sequence (Figure 24B-D), of Destination Vector pDEST4. This vector may also be referred to as pTrx-DEST4.

Figure 25 is a schematic depiction of the attR1 and attR2 sites (Figure 25A), the physical map (Figure 25B), and the nucleotide sequence (Figure 25C-D), of Destination Vector pDEST5. This vector may also be referred to as pSPORT(+)-DEST5.

Figure 26 is a schematic depiction of the attR1 and attR2 sites (Figure 26A), the physical map (Figure 26B), and the nucleotide sequence (Figure 26C-D), of Destination Vector pDEST6. This vector may also be referred to as pSPORT(-)-DEST6.

Figure 27 is a schematic depiction of the attR1 site, CMV promoter, and the physical map (Figure 27A), and the nucleotide sequence (Figure 27B-C), of Destination Vector pDEST7. This vector may also be referred to as pCMV-DEST7.

Figure 28 is a schematic depiction of the attR1 site, baculovirus polyhedrin promoter, and the physical map (Figure 28A), and the nucleotide sequence (Figure 28B-D), of Destination Vector pDEST8. This vector may also be referred to as pFastBac-DEST8.

Figure 29 is a schematic depiction of the attR1 site, Semliki Forest Virus promoter, and the physical map (Figure 29A), and the nucleotide sequence (Figure 29B-E), of Destination Vector pDEST9. This vector may also be referred to as pSFV-DEST9.

Figure 30 is a schematic depiction of the attR1 site, baculovirus polyhedrin promoter, His6 fusion domain, and the physical map (Figure 30A), and the nucleotide sequence (Figure 30B-D), of Destination Vector pDEST10. This vector may also be referred to as pFastBacHT-DEST10.

Figure 31 is a schematic depiction of the attR1 cassette containing a tetracycline-regulated CMV promoter and the physical map (Figure 31A), and the nucleotide sequence (Figure 31B-D), of Destination Vector pDEST11. This vector may also be referred to as pTet-DEST11.

Figure 32 is a schematic depiction of the attR1 site, the start of the mRNA of the CMV promoter, and the physical map (Figure 32A), and the nucleotide sequence (Figure 32B-D), of Destination Vector pDEST12.2. This vector may also be referred to as pCMVneo-DEST12, as pCMV-DEST12, or as pDEST12.

Figure 33 is a schematic depiction of the attR1 site, the λP_L promoter, and the physical map (Figure 33A), and the nucleotide sequence (Figure 33B-C), of Destination Vector pDEST13. This vector may also be referred to as p λP_L -DEST13.

Figure 34 is a schematic depiction of the attR1 site, the T7 promoter, and the physical map (Figure 34A), and the nucleotide sequence (Figure 34B-D), of Destination Vector pDEST14. This vector may also be referred to as pPT7-DEST14.

Figure 35 is a schematic depiction of the attR1 site, the T7 promoter, and the N-terminal GST fusion sequence, and the physical map (Figure 35A), and the nucleotide sequence (Figure 35B-D), of Destination Vector pDEST15. This vector may also be referred to as pT7 GST-DEST15.

Figure 36 is a schematic depiction of the attR1 site, the T7 promoter, and the N-terminal thioredoxin fusion sequence, and the physical map (Figure 36A), and the nucleotide sequence (Figure 36B-D), of Destination Vector pDEST16. This vector may also be referred to as pT7 Trx-DEST16.

Figure 37 is a schematic depiction of the attR1 site, the T7 promoter, and the N-terminal His6 fusion sequence, and the physical map (Figure 37A), and the

nucleotide sequence (Figure 37B-D), of Destination Vector pDEST17. This vector may also be referred to as pT7 His-DEST17.

5 **Figure 38** is a schematic depiction of the attR1 site and the p10 baculovirus promoter, and the physical map (Figure 38A), and the nucleotide sequence (Figure 38B-D), of Destination Vector pDEST18. This vector may also be referred to as pFBp10-DEST18.

10 **Figure 39** is a schematic depiction of the attR1 site, and the 39k baculovirus promoter, and the physical map (Figure 39A), and the nucleotide sequence (Figure 39B-D), of Destination Vector pDEST19. This vector may also be referred to as pFB39k-DEST19.

15 **Figure 40** is a schematic depiction of the attR1 site, the *polh* baculovirus promoter, and the N-terminal GST fusion sequence, and the physical map (Figure 40A), and the nucleotide sequence (Figure 40B-D), of Destination Vector pDEST20. This vector may also be referred to as pFB GST-DEST20.

20 **Figure 41** is a schematic depiction of a 2-hybrid vector with a DNA-binding domain, the attR1 site, and the ADH promoter, and the physical map (Figure 41A), and the nucleotide sequence (Figure 41B-E), of Destination Vector pDEST21. This vector may also be referred to as pDB Leu-DEST21.

25 **Figure 42** is a schematic depiction of a 2-hybrid vector with an activation domain, the attR1 site, and the ADH promoter, and the physical map (Figure 42A), and the nucleotide sequence (Figure 42B-D), of Destination Vector pDEST22. This vector may also be referred to as pPC86-DEST22.

30 **Figure 43** is a schematic depiction of the attR1 and attR2 sites, the T7 promoter, and the C-terminal His6 fusion sequence, and the physical map (Figure 43A), and the nucleotide sequence (Figure 43B-D), of Destination Vector pDEST23. This vector may also be referred to as pC-term-His6-DEST23.

Figure 44 is a schematic depiction of the attR1 and attR2 sites, the T7 promoter, and the C-terminal GST fusion sequence, and the physical map (Figure 44A), and the nucleotide sequence (Figure 44B-D), of Destination Vector pDEST24. This vector may also be referred to as pC-term-GST-DEST24.

Figure 45 is a schematic depiction of the attR1 and attR2 sites, the T7 promoter, and the C-terminal thioredoxin fusion sequence, and the physical map (Figure 45A), and the nucleotide sequence (Figure 45B-D), of Destination Vector pDEST25. This vector may also be referred to as pC-term-Trx-DEST25.

Figure 46 is a schematic depiction of the attR1 site, the CMV promoter, and an N-terminal His6 fusion sequence, and the physical map (Figure 46A), and the nucleotide sequence (Figure 46B-D), of Destination Vector pDEST26. This vector may also be referred to as pCMV-SPneo-His-DEST26.

Figure 47 is a schematic depiction of the attR1 site, the CMV promoter, and an N-terminal GST fusion sequence, and the physical map (Figure 47A), and the nucleotide sequence (Figure 47B-D), of Destination Vector pDEST27. This vector may also be referred to as pCMV-Spneo-GST-DEST27.

Figure 48 is a depiction of the physical map (Figure 48A), the cloning sites (Figure 48B), and the nucleotide sequence (Figure 48C-D), for the attB cloning vector plasmid pEXP501. This vector may also be referred to equivalently herein as pCMV•SPORT6, pCMVSPORT6, and pCMVSPORT6.

Figure 49 is a depiction of the physical map (Figure 49A), and the nucleotide sequence (Figure 49B-C), for the Donor plasmid pDONR201 which donates a kanamycin-resistant vector in the BP Reaction. This vector may also be referred to as pAttPkanr Donor Plasmid, or as pAttPkan Donor Plasmid

Figure 50 is a depiction of the physical map (Figure 50A), and the nucleotide sequence (Figure 50B-C), for the Donor plasmid pDONR202 which donates a kanamycin-resistant vector in the BP Reaction.

Figure 51 is a depiction of the physical map (Figure 51A), and the nucleotide sequence (Figure 51B-C), for the Donor plasmid pDONR203 which donates a kanamycin-resistant vector in the BP Reaction.

Figure 52 is a depiction of the physical map (Figure 52A), and the nucleotide sequence (Figure 52B-C), for the Donor plasmid pDONR204 which donates a kanamycin-resistant vector in the BP Reaction.

Figure 53 is a depiction of the physical map (Figure 53A), and the nucleotide sequence (Figure 53B-C), for the Donor plasmid pDONR205 which donates a tetracycline-resistant vector in the BP Reaction.

Figure 54 is a depiction of the physical map (Figure 54A), and the nucleotide sequence (Figure 54B-C), for the Donor plasmid pDONR206 which donates a gentamycin-resistant vector in the BP Reaction. This vector may also be referred to as pENTR22 attP Donor Plasmid, pAttPGenr Donor Plasmid, or pAttPgnt Donor Plasmid.

Figure 55 depicts the attB1 site, and the physical map, of an Entry Clone (pENTR7) of CAT subcloned into the Destination Vector pDEST2 (Figure 22).

Figure 56 depicts the DNA components of Reaction B of the one-tube BxP reaction described in Example 16, pEZC7102 and attB-tet-PCR.

Figure 57 is a physical map of the desired product of Reaction B of the one-tube BxP reaction described in Example 16, tetx7102.

Figure 58 is a physical map of the Destination Vector pEZC8402.

Figure 59 is a physical map of the expected tet^r subclone product, tetx8402, resulting from the LxR Reaction with tetx7102 (Figure 57) plus pEZC8402 (Figure 58).

Figure 60 is a schematic depiction of the bacteriophage lambda recombination pathways in *E. coli*.

Figure 61 is a schematic depiction of the DNA molecules participating in the LR Reaction. Two different co-integrates form during the LR Reaction (only one of which is shown here), depending on whether attL1 and attR1 or attL2 and attR2 are first to recombine. In one aspect, the invention provides directional cloning of a nucleic acid molecule of interest, since the recombination sites react with specificity (attL1 reacts with attR1; attL2 with attR2; attB1 with attP1; and attB2 with attP2). Thus, positioning of the sites allows construction of desired vectors having recombined fragments in the desired orientation.

Figure 62 is a depiction of native and fusion protein expression using the recombinational cloning methods and compositions of the invention. In the upper figure depicting native protein expression, all of the translational start signals are

included between the attB1 and attB2 sites; therefore, these signals must be present in the starting Entry Clone. The lower figure depicts fusion protein expression (here showing expression with both N-terminal and C-terminal fusion tags so that ribosomes read through attB1 and attB2 to create the fusion protein). Unlike native protein expression vectors, N-terminal fusion vectors have their translational start signals upstream of the attB1 site.

Figure 63 is a schematic depiction of three GATEWAY™ Cloning System cassettes. Three blunt-ended cassettes are depicted which convert standard expression vectors to Destination Vectors. Each of the depicted cassettes provides amino-terminal fusions in one of three possible reading frames, and each has a distinctive restriction cleavage site as shown.

Figure 64 shows the physical maps of plasmids containing three attR reading frame cassettes, pEZX15101 (reading frame A; Figure 64A), pEZX15102 (reading frame B; Figure 64B), and pEZX15103 (reading frame C; Figure 64C).

Figure 65 depicts the attB primers used for amplifying the tet^r and amp^r genes from pBR322 by the cloning methods of the invention.

Figure 66 is a table listing the results of recombinational cloning of the tet^r and amp^r PCR products made using the primers shown in Figure 65.

Figure 67 is a graph showing the effect of the number of guanines (G's) contained on the 5' end of the PCR primers on the cloning efficiency of PCR products. It is noted, however, that other nucleotides besides guanine (including A, T, C, U or combinations thereof) may be used as 5' extensions on the PCR primers to enhance cloning efficiency of PCR products.

Figure 68 is a graph showing a titration of various amounts of attP and attB reactants in the BxP reaction, and the effects on cloning efficiency of PCR products.

Figure 69 is a series of graphs showing the effects of various weights (Figure 69A) or moles (Figure 69B) of a 256 bp PCR product on formation of colonies, and on efficiency of cloning of the 256 bp PCR product into a Donor Vector (Figure 69C).

Figure 70 is a series of graphs showing the effects of various weights (Figure 70A) or moles (Figure 70B) of a 1 kb PCR product on formation of colonies, and on efficiency of cloning of the 1 kb PCR product into a Donor Vector (Figure 70C).

5 Figure 71 is a series of graphs showing the effects of various weights (Figure 71A) or moles (Figure 71B) of a 1.4 kb PCR product on formation of colonies, and on efficiency of cloning of the 1.4 kb PCR product into a Donor Vector (Figure 71C).

10 Figure 72 is a series of graphs showing the effects of various weights (Figure 72A) or moles (Figure 72B) of a 3.4 kb PCR product on formation of colonies, and on efficiency of cloning of the 3.4 kb PCR product into a Donor Vector (Figure 72C).

15 Figure 73 is a series of graphs showing the effects of various weights (Figure 73A) or moles (Figure 73B) of a 4.6 kb PCR product on formation of colonies, and on efficiency of cloning of the 4.6 kb PCR product into a Donor Vector (Figure 73C).

Figure 74 is photograph of an ethidium bromide-stained gel of a titration of a 6.9 kb PCR product in a BxP reaction.

20 Figure 75 is a graph showing the effects of various amounts of a 10.1 kb PCR product on formation of colonies upon cloning of the 10.1 kb PCR product into a Donor Vector.

Figure 76 is photograph of an ethidium bromide-stained gel of a titration of a 10.1 kb PCR product in a BxP reaction.

25 Figure 77 is a table summarizing the results of the PCR product cloning efficiency experiments depicted in Figures 69-74, for PCR fragments ranging in size from 0.256 kb to 6.9 kb.

30 Figure 78 is a depiction of the sequences at the ends of attR Cassettes. Sequences contributed by the Cm^r -ccdB cassette are shown, including the outer ends of the flanking attR sites (boxed). The staggered cleavage sites for Int are indicated in the boxed regions. Following recombination with an Entry Clone, only the outer sequences in attR sites contribute to the resulting attB sites in the

Expression Clone. The underlined sequences at both ends dictate the different reading frames (reading frames A, B, or C, with two alternative reading frame C cassettes depicted) for fusion proteins.

Figure 79 is a depiction of several different attR cassettes (in reading frames A, B, or C) which may provide fusion codons at the amino-terminus of the encoded protein.

Figure 80 illustrates the single-cutting restriction sites in an attR reading frame A cassette of the invention.

Figure 81 illustrates the single-cutting restriction sites in an attR reading frame B cassette of the invention.

Figure 82 illustrates the single-cutting restriction sites in two alternative attR reading frame C cassettes of the invention (Figures 82A and 82B) depicted in Figure 78.

Figure 83 shows the physical map (Figure 83A), and the nucleotide sequence (Figure 83B-C), for an attR reading frame C parent plasmid prfC Parent III, which contains an attR reading frame C cassette of the invention (alternative A in Figures 78 and 82).

Figure 84 is a physical map of plasmid pEYC1301.

Figure 85 is a physical map of plasmid pEYC1313.

Figure 86 is a physical map of plasmid pEYC14032.

Figure 87 is a physical map of plasmid pMAB58.

Figure 88 is a physical map of plasmid pMAB62.

Figure 89 is a depiction of a synthesis reaction using two pairs of homologous primers of the invention.

Figure 90 is a schematic depiction of the physical map (Figure 90A), and the nucleotide sequence (Figure 90B-D), of Destination Vector pDEST28.

Figure 91 is a schematic depiction of the physical map (Figure 91A), and the nucleotide sequence (Figure 91B-D), of Destination Vector pDEST29.

Figure 92 is a schematic depiction of the physical map (Figure 92A), and the nucleotide sequence (Figure 92B-D), of Destination Vector pDEST30.

Figure 93 is a schematic depiction of the physical map (Figure 93A), and the nucleotide sequence (Figure 93B-D), of Destination Vector pDEST31.

Figure 94 is a schematic depiction of the physical map (Figure 94A), and the nucleotide sequence (Figure 94B-E), of Destination Vector pDEST32.

5 **Figure 95** is a schematic depiction of the physical map (Figure 95A), and the nucleotide sequence (Figure 95B-D), of Destination Vector pDEST33.

Figure 96 is a schematic depiction of the physical map (Figure 96A), and the nucleotide sequence (Figure 96B-D), of Destination Vector pDEST34.

10 **Figure 97** is a depiction of the physical map (Figure 97A), and the nucleotide sequence (Figure 97B-C), for the Donor plasmid pDONR207 which donates a gentamycin-resistant vector in the BP Reaction.

Figure 98 is a schematic depiction of the physical map (Figure 98A), and the nucleotide sequence (Figure 98B-D), of the 2-hybrid vector pMAB85.

15 **Figure 99** is a schematic depiction of the physical map (Figure 99A), and the nucleotide sequence (Figure 99B-D), of the 2-hybrid vector pMAB86.

DETAILED DESCRIPTION OF THE INVENTION

20 *Definitions*

In the description that follows, a number of terms used in recombinant DNA technology are utilized extensively. In order to provide a clear and consistent understanding of the specification and claims, including the scope to be given such terms, the following definitions are provided.

25 **Byproduct:** is a daughter molecule (a new clone produced after the second recombination event during the recombinational cloning process) lacking the segment which is desired to be cloned or subcloned.

30 **Cointegrate:** is at least one recombination intermediate nucleic acid molecule of the present invention that contains both parental (starting) molecules. It will usually be linear. In some embodiments it can be circular. RNA and polypeptides may be expressed from cointegrates using an appropriate host cell strain, for example *E. coli* DB3.1 (particularly *E. coli* LIBRARY EFFICIENCY®

DB3.1™ Competent Cells), and selecting for both selection markers found on the cointegrate molecule.

Host: is any prokaryotic or eukaryotic organism that can be a recipient of the recombinational cloning Product, vector, or nucleic acid molecule of the invention. A "host," as the term is used herein, includes prokaryotic or eukaryotic organisms that can be genetically engineered. For examples of such hosts, see Maniatis *et al.*, *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York (1982).

Insert or Inserts: include the desired nucleic acid segment or a population of nucleic acid segments (segment *A* of Figure 1) which may be manipulated by the methods of the present invention. Thus, the terms Insert(s) are meant to include a particular nucleic acid (preferably DNA) segment or a population of segments. Such Insert(s) can comprise one or more nucleic acid molecules.

Insert Donor: is one of the two parental nucleic acid molecules (e.g. RNA or DNA) of the present invention which carries the Insert. The Insert Donor molecule comprises the Insert flanked on both sides with recombination sites. The Insert Donor can be linear or circular. In one embodiment of the invention, the Insert Donor is a circular DNA molecule and further comprises a cloning vector sequence outside of the recombination signals (see Figure 1). When a population of Inserts or population of nucleic acid segments are used to make the Insert Donor, a population of Insert Donors results and may be used in accordance with the invention. Examples of such Insert Donor molecules are GATEWAY™ Entry Vectors, which include but are not limited to those Entry Vectors depicted in Figures 10-20, as well as other vectors comprising a gene of interest flanked by one or more *attL* sites (e.g., *attL1*, *attL2*, etc.), or by one or more *attB* sites (e.g., *attB1*, *attB2*, etc.) for the production of library clones.

Product: is one of the desired daughter molecules comprising the *A* and *D* sequences which is produced after the second recombination event during the recombinational cloning process (see Figure 1). The Product contains the nucleic acid which was to be cloned or subcloned. In accordance with the invention, when a population of Insert Donors are used, the resulting population of Product

molecules will contain all or a portion of the population of Inserts of the Insert Donors and preferably will contain a representative population of the original molecules of the Insert Donors.

Promoter: is a DNA sequence generally described as the 5'-region of a gene, located proximal to the start codon. The transcription of an adjacent DNA segment is initiated at the promoter region. A repressible promoter's rate of transcription decreases in response to a repressing agent. An inducible promoter's rate of transcription increases in response to an inducing agent. A constitutive promoter's rate of transcription is not specifically regulated, though it can vary under the influence of general metabolic conditions.

Recognition sequence: Recognition sequences are particular sequences which a protein, chemical compound, DNA, or RNA molecule (*e.g.*, restriction endonuclease, a modification methylase, or a recombinase) recognizes and binds. In the present invention, a recognition sequence will usually refer to a recombination site. For example, the recognition sequence for Cre recombinase is *loxP* which is a 34 base pair sequence comprised of two 13 base pair inverted repeats (serving as the recombinase binding sites) flanking an 8 base pair core sequence. See Figure 1 of Sauer, B., *Current Opinion in Biotechnology* 5:521-527 (1994). Other examples of recognition sequences are the *attB*, *attP*, *attL*, and *attR* sequences which are recognized by the recombinase enzyme λ Integrase. *attB* is an approximately 25 base pair sequence containing two 9 base pair core-type Int binding sites and a 7 base pair overlap region. *attP* is an approximately 240 base pair sequence containing core-type Int binding sites and arm-type Int binding sites as well as sites for auxiliary proteins integration host factor (IHF), FIS and excisionase (Xis). See Landy, *Current Opinion in Biotechnology* 3:699-707 (1993). Such sites may also be engineered according to the present invention to enhance production of products in the methods of the invention. When such engineered sites lack the P1 or H1 domains to make the recombination reactions irreversible (*e.g.*, *attR* or *attP*), such sites may be designated *attR'* or *attP'* to show that the domains of these sites have been modified in some way.

Recombination proteins: include excisive or integrative proteins, enzymes, co-factors or associated proteins that are involved in recombination reactions involving one or more recombination sites, which may be wild-type proteins (See Landy, *Current Opinion in Biotechnology* 3:699-707 (1993)), or mutants, derivatives (e.g., fusion proteins containing the recombination protein sequences or fragments thereof), fragments, and variants thereof.

Recombination site: is a recognition sequence on a DNA molecule participating in an integration/recombination reaction by the recombinational cloning methods of the invention. Recombination sites are discrete sections or segments of DNA on the participating nucleic acid molecules that are recognized and bound by a site-specific recombination protein during the initial stages of integration or recombination. For example, the recombination site for Cre recombinase is *loxP* which is a 34 base pair sequence comprised of two 13 base pair inverted repeats (serving as the recombinase binding sites) flanking an 8 base pair core sequence. See Figure 1 of Sauer, B., *Curr. Opin. Biotech.* 5:521-527 (1994). Other examples of recognition sequences include the *attB*, *attP*, *attL*, and *attR* sequences described herein, and mutants, fragments, variants and derivatives thereof, which are recognized by the recombination protein λ Int and by the auxiliary proteins integration host factor (IHF), FIS and excisionase (Xis). See Landy, *Curr. Opin. Biotech.* 3:699-707 (1993).

Recombinational Cloning: is a method described herein, whereby segments of nucleic acid molecules or populations of such molecules are exchanged, inserted, replaced, substituted or modified, *in vitro* or *in vivo*. By "*in vitro*" and "*in vivo*" herein is meant recombinational cloning that is carried out outside of host cells (e.g., in cell-free systems) or inside of host cells (e.g., using recombination proteins expressed by host cells), respectively.

Repression cassette: is a nucleic acid segment that contains a repressor or a Selectable marker present in the subcloning vector.

Selectable marker: is a DNA segment that allows one to select for or against a molecule (e.g., a replicon) or a cell that contains it, often under particular conditions. These markers can encode an activity, such as, but not limited to,

production of RNA, peptide, or protein, or can provide a binding site for RNA, peptides, proteins, inorganic and organic compounds or compositions and the like. Examples of Selectable markers include but are not limited to: (1) DNA segments that encode products which provide resistance against otherwise toxic compounds (e.g., antibiotics); (2) DNA segments that encode products which are otherwise lacking in the recipient cell (e.g., tRNA genes, auxotrophic markers); (3) DNA segments that encode products which suppress the activity of a gene product; (4) DNA segments that encode products which can be readily identified (e.g., phenotypic markers such as β -galactosidase, green fluorescent protein (GFP), and cell surface proteins); (5) DNA segments that bind products which are otherwise detrimental to cell survival and/or function; (6) DNA segments that otherwise inhibit the activity of any of the DNA segments described in Nos. 1-5 above (e.g., antisense oligonucleotides); (7) DNA segments that bind products that modify a substrate (e.g. restriction endonucleases); (8) DNA segments that can be used to isolate or identify a desired molecule (e.g. specific protein binding sites); (9) DNA segments that encode a specific nucleotide sequence which can be otherwise non-functional (e.g., for PCR amplification of subpopulations of molecules); (10) DNA segments, which when absent, directly or indirectly confer resistance or sensitivity to particular compounds; (11) DNA segments that encode products which are toxic in recipient cells; (12) DNA segments that inhibit replication, partition or heritability of nucleic acid molecules that contain them; and/or (13) DNA segments that encode conditional replication functions, e.g., replication in certain hosts or host cell strains or under certain environmental conditions (e.g., temperature, nutritional conditions, etc.).

Selection scheme: is any method which allows selection, enrichment, or identification of a desired Product or Product(s) from a mixture containing an Entry Clone or Vector, a Destination Vector, a Donor Vector, an Expression Clone or Vector, any intermediates (e.g. a Cointegrate or a replicon), and/or Byproducts. The selection schemes of one preferred embodiment have at least two components that are either linked or unlinked during recombinational cloning. One component is a Selectable marker. The other component controls the expression *in vitro* or *in vivo* of the Selectable marker, or survival of the cell (or

the nucleic acid molecule, *e.g.*, a replicon) harboring the plasmid carrying the Selectable marker. Generally, this controlling element will be a repressor or inducer of the Selectable marker, but other means for controlling expression or activity of the Selectable marker can be used. Whether a repressor or activator is used will depend on whether the marker is for a positive or negative selection, and the exact arrangement of the various DNA segments, as will be readily apparent to those skilled in the art. A preferred requirement is that the selection scheme results in selection of or enrichment for only one or more desired Products. As defined herein, selecting for a DNA molecule includes (a) selecting or enriching for the presence of the desired DNA molecule, and (b) selecting or enriching against the presence of DNA molecules that are not the desired DNA molecule.

In one embodiment, the selection schemes (which can be carried out in reverse) will take one of three forms, which will be discussed in terms of Figure 1. The first, exemplified herein with a Selectable marker and a repressor therefore, selects for molecules having segment *D* and lacking segment *C*. The second selects against molecules having segment *C* and for molecules having segment *D*. Possible embodiments of the second form would have a DNA segment carrying a gene toxic to cells into which the *in vitro* reaction products are to be introduced. A toxic gene can be a DNA that is expressed as a toxic gene product (a toxic protein or RNA), or can be toxic in and of itself. (In the latter case, the toxic gene is understood to carry its classical definition of "heritable trait".)

Examples of such toxic gene products are well known in the art, and include, but are not limited to, restriction endonucleases (*e.g.*, *DpnI*), apoptosis-related genes (*e.g.* ASK1 or members of the *bcl-2/ced-9* family), retroviral genes including those of the human immunodeficiency virus (HIV), defensins such as NP-1, inverted repeats or paired palindromic DNA sequences, bacteriophage lytic genes such as those from Φ X174 or bacteriophage T4; antibiotic sensitivity genes such as *rpsL*, antimicrobial sensitivity genes such as *pheS*, plasmid killer genes, eukaryotic transcriptional vector genes that produce a gene product toxic to bacteria, such as GATA-1, and genes that kill hosts in the absence of a suppressing function, *e.g.*, *kicB*, *ccdB*, Φ X174 *E* (Liu, Q. *et al.*, *Curr. Biol.*

8:1300-1309 (1998)), and other genes that negatively affect replicon stability and/or replication. A toxic gene can alternatively be selectable *in vitro*, e.g., a restriction site.

5 Many genes coding for restriction endonucleases operably linked to inducible promoters are known, and may be used in the present invention. See, e.g. U.S. Patent Nos. 4,960,707 (*DpnI* and *DpnII*); 5,000,333, 5,082,784 and 5,192,675 (*KpnI*); 5,147,800 (*NgoAIII* and *NgoAI*); 5,179,015 (*FspI* and *HaeIII*); 5,200,333 (*HaeII* and *TaqI*); 5,248,605 (*HpaII*); 5,312,746 (*ClaI*); 5,231,021 and 5,304,480 (*XhoI* and *XhoII*); 5,334,526 (*AluI*); 5,470,740 (*NsiI*); 5,534,428
10 (*SstI/SacI*); 5,202,248 (*NcoI*); 5,139,942 (*NdeI*); and 5,098,839 (*PacI*). See also Wilson, G.G., *Nucl. Acids Res.* 19:2539-2566 (1991); and Lunnen, K.D., *et al.*, *Gene* 74:25-32 (1988).

In the second form, segment *D* carries a Selectable marker. The toxic gene would eliminate transformants harboring the Vector Donor, Cointegrate, and
15 Byproduct molecules, while the Selectable marker can be used to select for cells containing the Product and against cells harboring only the Insert Donor.

The third form selects for cells that have both segments *A* and *D* in *cis* on the same molecule, but not for cells that have both segments in *trans* on different molecules. This could be embodied by a Selectable marker that is split into two
20 inactive fragments, one each on segments *A* and *D*.

The fragments are so arranged relative to the recombination sites that when the segments are brought together by the recombination event, they reconstitute a functional Selectable marker. For example, the recombinational event can link a promoter with a structural nucleic acid molecule (e.g., a gene),
25 can link two fragments of a structural nucleic acid molecule, or can link nucleic acid molecules that encode a heterodimeric gene product needed for survival, or can link portions of a replicon.

Site-specific recombinase: is a type of recombinase which typically has at least the following four activities (or combinations thereof): (1) recognition of
30 one or two specific nucleic acid sequences; (2) cleavage of said sequence or sequences; (3) topoisomerase activity involved in strand exchange; and (4) ligase

activity to reseal the cleaved strands of nucleic acid. See Sauer, B., *Current Opinions in Biotechnology* 5:521-527 (1994). Conservative site-specific recombination is distinguished from homologous recombination and transposition by a high degree of sequence specificity for both partners. The strand exchange mechanism involves the cleavage and rejoining of specific DNA sequences in the absence of DNA synthesis (Landy, A. (1989) *Ann. Rev. Biochem.* 58:913-949).

Subcloning vector: is a cloning vector comprising a circular or linear nucleic acid molecule which includes preferably an appropriate replicon. In the present invention, the subcloning vector (segment *D* in Figure 1) can also contain functional and/or regulatory elements that are desired to be incorporated into the final product to act upon or with the cloned DNA Insert (segment *A* in Figure 1). The subcloning vector can also contain a Selectable marker (preferably DNA).

Vector: is a nucleic acid molecule (preferably DNA) that provides a useful biological or biochemical property to an Insert. Examples include plasmids, phages, autonomously replicating sequences (ARS), centromeres, and other sequences which are able to replicate or be replicated *in vitro* or in a host cell, or to convey a desired nucleic acid segment to a desired location within a host cell. A Vector can have one or more restriction endonuclease recognition sites at which the sequences can be cut in a determinable fashion without loss of an essential biological function of the vector, and into which a nucleic acid fragment can be spliced in order to bring about its replication and cloning. Vectors can further provide primer sites, *e.g.*, for PCR, transcriptional and/or translational initiation and/or regulation sites, recombinational signals, replicons, Selectable markers, *etc.* Clearly, methods of inserting a desired nucleic acid fragment which do not require the use of homologous recombination, transpositions or restriction enzymes (such as, but not limited to, UDG cloning of PCR fragments (U.S. Patent No. 5,334,575, entirely incorporated herein by reference), T:A cloning, and the like) can also be applied to clone a fragment into a cloning vector to be used according to the present invention. The cloning vector can further contain one or more selectable markers suitable for use in the identification of cells transformed with the cloning vector.

Vector Donor: is one of the two parental nucleic acid molecules (e.g. RNA or DNA) of the present invention which carries the DNA segments comprising the DNA vector which is to become part of the desired Product. The Vector Donor comprises a subcloning vector *D* (or it can be called the cloning vector if the Insert Donor does not already contain a cloning vector (e.g., for PCR fragments containing *attB* sites; see below)) and a segment *C* flanked by recombination sites (see Figure 1). Segments *C* and/or *D* can contain elements that contribute to selection for the desired Product daughter molecule, as described above for selection schemes. The recombination signals can be the same or different, and can be acted upon by the same or different recombinases. In addition, the Vector Donor can be linear or circular. Examples of such Vector Donor molecules include GATEWAY™ Destination Vectors, which include but are not limited to those Destination Vectors depicted in Figures 21-47 and 90-96.

Primer: refers to a single stranded or double stranded oligonucleotide that is extended by covalent bonding of nucleotide monomers during amplification or polymerization of a nucleic acid molecule (e.g. a DNA molecule). In a preferred aspect, a primer comprises one or more recombination sites or portions of such recombination sites. Portions of recombination sites comprise at least 2 bases (or basepairs, abbreviated herein as "bp"), at least 5-200 bases, at least 10-100 bases, at least 15-75 bases, at least 15-50 bases, at least 15-25 bases, or at least 16-25 bases, of the recombination sites of interest, as described in further detail below and in the Examples. When using portions of recombination sites, the missing portion of the recombination site may be provided as a template by the newly synthesized nucleic acid molecule. Such recombination sites may be located within and/or at one or both termini of the primer. Preferably, additional sequences are added to the primer adjacent to the recombination site(s) to enhance or improve recombination and/or to stabilize the recombination site during recombination. Such stabilization sequences may be any sequences (preferably G/C rich sequences) of any length. Preferably, such sequences range in size from 1 to about 1000 bases, 1 to about 500 bases, and 1 to about 100 bases, 1 to about 60 bases, 1 to about 25, 1 to about 10, 2 to about 10 and preferably about 4 bases.

Preferably, such sequences are greater than 1 base in length and preferably greater than 2 bases in length.

Template: refers to double stranded or single stranded nucleic acid molecules which are to be amplified, synthesized or sequenced. In the case of double stranded molecules, denaturation of its strands to form a first and a second strand is preferably performed before these molecules will be amplified, synthesized or sequenced, or the double stranded molecule may be used directly as a template. For single stranded templates, a primer complementary to a portion of the template is hybridized under appropriate conditions and one or more polypeptides having polymerase activity (e.g. DNA polymerases and/or reverse transcriptases) may then synthesize a nucleic acid molecule complementary to all or a portion of said template. Alternatively, for double stranded templates, one or more promoters may be used in combination with one or more polymerases to make nucleic acid molecules complementary to all or a portion of the template. The newly synthesized molecules, according to the invention, may be equal or shorter in length than the original template. Additionally, a population of nucleic acid templates may be used during synthesis or amplification to produce a population of nucleic acid molecules typically representative of the original template population.

Adapter: is an oligonucleotide or nucleic acid fragment or segment (preferably DNA) which comprises one or more recombination sites (or portions of such recombination sites) which in accordance with the invention can be added to a circular or linear Insert Donor molecule as well as other nucleic acid molecules described herein. When using portions of recombination sites, the missing portion may be provided by the Insert Donor molecule. Such adapters may be added at any location within a circular or linear molecule, although the adapters are preferably added at or near one or both termini of a linear molecule. Preferably, adapters are positioned to be located on both sides (flanking) a particular nucleic acid molecule of interest. In accordance with the invention, adapters may be added to nucleic acid molecules of interest by standard recombinant techniques (e.g. restriction digest and ligation). For example, adapters may be added to a circular molecule by first digesting the molecule with

an appropriate restriction enzyme, adding the adapter at the cleavage site and reforming the circular molecule which contains the adapter(s) at the site of cleavage. In other aspects, adapters may be added by homologous recombination, by integration of RNA molecules, and the like. Alternatively, adapters may be ligated directly to one or more and preferably both termini of a linear molecule thereby resulting in linear molecule(s) having adapters at one or both termini. In one aspect of the invention, adapters may be added to a population of linear molecules, (e.g. a cDNA library or genomic DNA which has been cleaved or digested) to form a population of linear molecules containing adapters at one and preferably both termini of all or substantial portion of said population.

Adapter-Primer: is primer molecule which comprises one or more recombination sites (or portions of such recombination sites) which in accordance with the invention can be added to a circular or linear nucleic acid molecule described herein. When using portions of recombination sites, the missing portion may be provided by a nucleic acid molecule (*e.g.*, an adapter) of the invention. Such adapter-primers may be added at any location within a circular or linear molecule, although the adapter-primers are preferably added at or near one or both termini of a linear molecule. Examples of such adapter-primers and the use thereof in accordance with the methods of the invention are shown in Example 25 herein. Such adapter-primers may be used to add one or more recombination sites or portions thereof to circular or linear nucleic acid molecules in a variety of contexts and by a variety of techniques, including but not limited to amplification (*e.g.*, PCR), ligation (*e.g.*, enzymatic or chemical/synthetic ligation), recombination (*e.g.*, homologous or non-homologous (illegitimate) recombination) and the like.

Library: refers to a collection of nucleic acid molecules (circular or linear). In one embodiment, a library may comprise a plurality (*i.e.*, two or more) of DNA molecules, which may or may not be from a common source organism, organ, tissue, or cell. In another embodiment, a library is representative of all or a portion or a significant portion of the DNA content of an organism (a "genomic" library), or a set of nucleic acid molecules representative of all or a portion or a significant portion of the expressed nucleic acid molecules (a cDNA library) in a

cell, tissue, organ or organism. A library may also comprise random sequences made by *de novo* synthesis, mutagenesis of one or more sequences and the like. Such libraries may or may not be contained in one or more vectors.

Amplification: refers to any *in vitro* method for increasing a number of copies of a nucleotide sequence with the use of a polymerase. Nucleic acid amplification results in the incorporation of nucleotides into a DNA and/or RNA molecule or primer thereby forming a new molecule complementary to a template. The formed nucleic acid molecule and its template can be used as templates to synthesize additional nucleic acid molecules. As used herein, one amplification reaction may consist of many rounds of replication. DNA amplification reactions include, for example, polymerase chain reaction (PCR). One PCR reaction may consist of 5-100 "cycles" of denaturation and synthesis of a DNA molecule.

Oligonucleotide: refers to a synthetic or natural molecule comprising a covalently linked sequence of nucleotides which are joined by a phosphodiester bond between the 3' position of the deoxyribose or ribose of one nucleotide and the 5' position of the deoxyribose or ribose of the adjacent nucleotide. This term may be used interchangeably herein with the terms "nucleic acid molecule" and "polynucleotide," without any of these terms necessarily indicating any particular length of the nucleic acid molecule to which the term specifically refers.

Nucleotide: refers to a base-sugar-phosphate combination. Nucleotides are monomeric units of a nucleic acid molecule (DNA and RNA). The term nucleotide includes ribonucleoside triphosphates ATP, UTP, CTG, GTP and deoxyribonucleoside triphosphates such as dATP, dCTP, dITP, dUTP, dGTP, dTTP, or derivatives thereof. Such derivatives include, for example, [α S]dATP, 7-deaza-dGTP and 7-deaza-dATP. The term nucleotide as used herein also refers to dideoxyribonucleoside triphosphates (ddNTPs) and their derivatives. Illustrated examples of dideoxyribonucleoside triphosphates include, but are not limited to, ddATP, ddCTP, ddGTP, ddITP, and ddTTP. According to the present invention, a "nucleotide" may be unlabeled or detectably labeled by well known techniques. Detectable labels include, for example, radioactive isotopes, fluorescent labels, chemiluminescent labels, bioluminescent labels and enzyme labels.

Hybridization: The terms "hybridization" and "hybridizing" refers to base pairing of two complementary single-stranded nucleic acid molecules (RNA and/or DNA) to give a double stranded molecule. As used herein, two nucleic acid molecules may be hybridized, although the base pairing is not completely complementary. Accordingly, mismatched bases do not prevent hybridization of two nucleic acid molecules provided that appropriate conditions, well known in the art, are used. In some aspects, hybridization is said to be under "stringent conditions." By "stringent conditions" as used herein is meant overnight incubation at 42°C in a solution comprising: 50% formamide, 5x SSC (150 mM NaCl, 15mM trisodium citrate), 50 mM sodium phosphate (pH 7.6), 5x Denhardt's solution, 10% dextran sulfate, and 20 g/ml denatured, sheared salmon sperm DNA, followed by washing the filters in 0.1x SSC at about 65°C.

Other terms used in the fields of recombinant DNA technology and molecular and cell biology as used herein will be generally understood by one of ordinary skill in the applicable arts.

Overview

Two reactions constitute the recombinational cloning system of the present invention, referred to herein as the "GATEWAY™ Cloning System," as depicted generally in Figure 1. The first of these reactions, the **LR Reaction** (Figure 2), which may also be referred to interchangeably herein as the **Destination Reaction**, is the main pathway of this system. The LR Reaction is a recombination reaction between an Entry vector or clone and a Destination Vector, mediated by a cocktail of recombination proteins such as the GATEWAY™ LR Clonase™ Enzyme Mix described herein. This reaction transfers nucleic acid molecules of interest (which may be genes, cDNAs, cDNA libraries, or fragments thereof) from the Entry Clone to an Expression Vector, to create an Expression Clone.

The sites labeled L, R, B, and P are respectively the attL, attR, attB, and attP recombination sites for the bacteriophage λ recombination proteins that constitute the Clonase cocktail (referred to herein variously as "Clonase" or

“GATEWAY™ LR Clonase™ Enzyme Mix” (for recombination protein mixtures mediating attL x attR recombination reactions, as described herein) or “GATEWAY™ BP Clonase™ Enzyme Mix” (for recombination protein mixtures mediating attB x attP recombination reactions, as described herein)). The
5 Recombinational Cloning reactions are equivalent to concerted, highly specific, cutting and ligation reactions. Viewed in this way, the recombination proteins cut to the left and right of the nucleic acid molecule of interest in the Entry Clone and ligate it into the Destination vector, creating a new Expression Clone.

The nucleic acid molecule of interest in an Expression Clone is flanked by
10 the small attB1 and attB2 sites. The orientation and reading frame of the nucleic acid molecule of interest are maintained throughout the subcloning, because attL1 reacts only with attR1, and attL2 reacts only with attR2. Likewise, attB1 reacts only with attP1, and attB2 reacts only with attP2. Thus, the invention also relates to methods of controlled or directional cloning using the recombination sites of the
15 invention (or portions thereof), including variants, fragments, mutants and derivatives thereof which may have altered or enhanced specificity. The invention also relates more generally to any number of recombination site partners or pairs (where each recombination site is specific for and interacts with its corresponding recombination site). Such recombination sites are preferably made by mutating or
20 modifying the recombination site to provide any number of necessary specificities (e.g., attB1-10, attP1-10, attL1-10, attR1-10, etc.), non-limiting examples of which are described in detail in the Examples herein.

When an aliquot from the recombination reaction is transformed into host cells (e.g., *E. coli*) and spread on plates containing an appropriate selection agent,
25 e.g., an antibiotic such as ampicillin with or without methicillin, cells that take up the desired clone form colonies. The unreacted Destination Vector does not give ampicillin-resistant colonies, even though it carries the ampicillin-resistance gene, because it contains a toxic gene, e.g., *ccdB*. Thus selection for ampicillin resistance selects for *E. coli* cells that carry the desired product, which usually
30 comprise >90% of the colonies on the ampicillin plate.

To participate in the Recombinational (or “GATEWAY™”) Cloning Reaction, a nucleic acid molecule of interest first may be cloned into an Entry

Vector, creating an Entry Clone. Multiple options are available for creating Entry Clones, including: cloning of PCR sequences with terminal attB recombination sites into Entry Vectors; using the GATEWAY™ Cloning System recombination reaction; transfer of genes from libraries prepared in GATEWAY™ Cloning System vectors by recombination into Entry Vectors; and cloning of restriction enzyme-generated fragments and PCR fragments into Entry Vectors by standard recombinant DNA methods. These approaches are discussed in further detail herein.

A key advantage of the GATEWAY™ Cloning System is that a nucleic acid molecule of interest (or even a population of nucleic acid molecules of interest) present as an Entry Clone can be subcloned in parallel into one or more Destination Vectors in a simple reactions for anywhere from about 30 seconds to about 60 minutes (preferably about 1-60 minutes, about 1-45 minutes, about 1-30 minutes, about 2-60 minutes, about 2-45 minutes, about 2-30 minutes, about 1-2 minutes, about 30-60 minutes, about 45-60 minutes, or about 30-45 minutes). Longer reaction times (*e.g.*, 2-24 hours, or overnight) may increase recombination efficiency, particularly where larger nucleic acid molecules are used, as described in the Examples herein. Moreover, a high percentage of the colonies obtained carry the desired Expression Clone. This process is illustrated schematically in Figure 3, which shows an advantage of the invention in which the molecule of interest can be moved simultaneously or separately into multiple Destination Vectors. In the LR Reaction, one or both of the nucleic acid molecules to be recombined may have any topology (*e.g.*, linear, relaxed circular, nicked circular, supercoiled, etc.), although one or both are preferably linear.

The second major pathway of the GATEWAY™ Cloning System is the **BP Reaction** (Figure 4), which may also be referred to interchangeably herein as the **Entry Reaction** or the **Gateward Reaction**. The BP Reaction may recombine an Expression Clone with a Donor Plasmid (the counterpart of the byproduct in Figure 2). This reaction transfers the nucleic acid molecule of interest (which may have any of a variety of topologies, including linear, coiled, supercoiled, etc.) in the Expression Clone into an Entry Vector, to produce a new Entry Clone. Once this nucleic acid molecule of interest is cloned into an Entry

Vector, it can be transferred into new Expression Vectors, through the LR Reaction as described above. In the BP Reaction, one or both of the nucleic acid molecules to be recombined may have any topology (*e.g.*, linear, relaxed circular, nicked circular, supercoiled, etc.), although one or both are preferably linear.

5 A useful variation of the BP Reaction permits rapid cloning and expression of products of amplification (*e.g.*, PCR) or nucleic acid synthesis. Amplification (*e.g.*, PCR) products synthesized with primers containing terminal 25 bp attB sites serve as efficient substrates for the Gateway Cloning reaction. Such amplification products may be recombined with a Donor Vector to produce an Entry Clone (see
10 Figure 7). The result is an Entry Clone containing the amplification fragment. Such Entry Clones can then be recombined with Destination Vectors -- through the LR Reaction -- to yield Expression Clones of the PCR product.

Additional details of the LR Reaction are shown in Figure 5A. The GATEWAY™ LR Clonase™ Enzyme Mix that mediates this reaction contains
15 lambda recombination proteins Int (Integrase), Xis (Excisionase), and IHF (Integration Host Factor). In contrast, the GATEWAY™ BP Clonase™ Enzyme Mix, which mediates the BP Reaction (Figure 5B), comprises Int and IHF alone.

The recombination (att) sites of each vector comprise two distinct segments, donated by the parental vectors. The staggered lines dividing the two
20 portions of each att site, depicted in Figures 5A and 5B, represent the seven-base staggered cut produced by Int during the recombination reactions. This structure is seen in greater detail in Figure 6, which displays the attB recombination sequences of an Expression Clone, generated by recombination between the attL1 and attL2 sites of an Entry Clone and the attR1 and attR2 sites of a Destination
25 Vector.

The nucleic acid molecule of interest in the Expression Clone is flanked by attB sites: attB1 to the left (amino terminus) and attB2 to the right (carboxy terminus). The bases in attB1 to the left of the seven-base staggered cut produced by Int are derived from the Destination vector, and the bases to the right of the
30 staggered cut are derived from the Entry Vector (see Figure 6). Note that the sequence is displayed in triplets corresponding to an open reading frame. If the reading frame of the nucleic acid molecule of interest cloned in the Entry Vector

is in phase with the reading frame shown for attB1, amino-terminal protein fusions can be made between the nucleic acid molecule of interest and any GATEWAY™ Cloning System Destination Vector encoding an amino-terminal fusion domain. Entry Vectors and Destination Vectors that enable cloning in all three reading frames are described in more detail herein, particularly in the Examples.

The LR Reaction allows the transfer of a desired nucleic acid molecule of interest into new Expression Vectors by recombining a Entry Clone with various Destination Vectors. To participate in the LR or Destination Reaction, however, a nucleic acid molecule of interest preferably is first converted to a Entry Clone. Entry Clones can be made in a number of ways, as shown in Figure 7.

One approach is to clone the nucleic acid molecule of interest into one or more of the Entry Vectors, using standard recombinant DNA methods, with restriction enzymes and ligase. The starting DNA fragment can be generated by restriction enzyme digestion or as a PCR product. The fragment is cloned between the attL1 and attL2 recombination sites in the Entry Vector. Note that a toxic or "death" gene (*e.g.*, *ccdB*), provided to minimize background colonies from incompletely digested Entry Vector, must be excised and replaced by the nucleic acid molecule of interest.

A second approach to making an Entry Clone (Figure 7) is to make a library (genomic or cDNA) in an Entry Vector, as described in detail herein. Such libraries may then be transferred into Destination Vectors for expression screening, for example in appropriate host cells such as yeast cells or mammalian cells.

A third approach to making Entry Clones (Figure 7) is to use Expression Clones obtained from cDNA molecules or libraries prepared in Expression Vectors. Such cDNAs or libraries, flanked by attB sites, can be introduced into a Entry Vector by recombination with a Donor Vector via the BP Reaction. If desired, an entire Expression Clone library can be transferred into the Entry Vector through the BP Reaction. Expression Clone cDNA libraries may also be constructed in a variety of prokaryotic and eukaryotic GATEWAY™-modified vectors (*e.g.*, the pEXP501 Expression Vector (see Figure 48), and 2-hybrid and

attB library vectors), as described in detail herein, particularly in the Examples below.

A fourth, and potentially most versatile, approach to making an Entry Clone (Figure 7) is to introduce a sequence for a nucleic acid molecule of interest into an Entry Vector by amplification (*e.g.*, PCR) fragment cloning. This method is diagramed in Figure 8. The DNA sequence first is amplified (for example, with PCR) as outlined in detail below and in the Examples herein, using primers containing one or more bp, two or more bp, three or more bp, four or more bp, five or more bp, preferably six or more bp, more preferably 6-25 bp (particularly 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24 or 25) bp of the attB nucleotide sequences (such as, but not limited to, those depicted in Figure 9), and optionally one or more, two or more, three or more, four or more, and most preferably four or five or more additional terminal nucleotide bases which preferably are guanines. The PCR product then may be converted to a Entry Clone by performing a BP Reaction, in which the attB-PCR product recombines with a Donor Vector containing one or more attP sites. Details of this approach and protocols for PCR fragment subcloning are provided in Examples 8 and 21-25.

A variety of Entry Clones may be produced by these methods, providing a wide array of cloning options; a number of specific Entry Vectors are also available commercially from Life Technologies, Inc. (Rockville, MD). The Examples herein provide a more in-depth description of selected Entry Vectors and details of their cloning sites. Choosing the optimal Entry Vector for a particular application is discussed in Example 4.

Entry Vectors and Destination Vectors should be constructed so that the amino-terminal region of a nucleic acid molecule of interest (*e.g.*, a gene, cDNA library or insert, or fragment thereof) will be positioned next to the attL1 site. Entry Vectors preferably contain the *rrnB* transcriptional terminator upstream of the attL1 site. This sequence ensures that expression of cloned nucleic acid molecules of interest is reliably "off" in *E. coli*, so that even toxic genes can be successfully cloned. Thus, Entry Clones may be designed to be transcriptionally silent. Note also that Entry Vectors, and hence Entry Clones, may contain the kanamycin antibiotic resistance (*kan^r*) gene to facilitate selection of host cells

containing Entry Clones after transformation. In certain applications, however, Entry Clones may contain other selection markers, including but not limited to a gentamycin resistance (*gen^r*) or tetracycline resistance (*tet^r*) gene, to facilitate selection of host cells containing Entry Clones after transformation.

5 Once a nucleic acid molecule of interest has been cloned into an Entry Vector, it may be moved into a Destination Vector. The upper right portion of Figure 5A shows a schematic of a Destination Vector. The thick arrow represents some function (often transcription or translation) that will act on the nucleic acid molecule of interest in the clone. During the recombination reaction, the region
10 between the *attR1* and *attR2* sites, including a toxic or "death" gene (*e.g.*, *ccdB*), is replaced by the DNA segment from the Entry Clone. Selection for recombinants that have acquired the ampicillin resistance (*amp^r*) gene (carried on the Destination Vector) and that have also lost the death gene ensures that a high percentage (usually >90%) of the resulting colonies will contain the correct insert.

15 To move a nucleic acid molecule of interest into a Destination Vector, the Destination Vector is mixed with the Entry Clone comprising the desired nucleic acid molecule of interest, a cocktail of recombination proteins (*e.g.*, GATEWAY™ LR Clonase™ Enzyme Mix) is added, the mixture is incubated (preferably at about 25°C for about 60 minutes, or longer under certain
20 circumstances, *e.g.* for transfer of large nucleic acid molecules, as described below) and any standard host cell (including bacterial cells such as *E. coli*; animal cells such as insect cells, mammalian cells, nematode cells and the like; plant cells; and yeast cells) strain is transformed with the reaction mixture. The host cell used will be determined by the desired selection (*e.g.*, *E. coli* DB3.1, available
25 commercially from Life Technologies, Inc., allows survival of clones containing the *ccdB* death gene, and thus can be used to select for cointegrate molecules -- *i.e.*, molecules that are hybrids between the Entry Clone and Destination Vector). The Examples below provide further details and protocols for use of Entry and Destination Vectors in transferring nucleic acid molecules of interest and
30 expressing RNAs or polypeptides encoded by these nucleic acid molecules in a variety of host cells.

The cloning system of the invention therefore offers multiple advantages:

- Once a nucleic acid molecule of interest is cloned into the GATEWAY™ Cloning System, it can be moved into and out of other vectors with complete fidelity of reading frame and orientation. That is, since the reactions proceed whereby attL1 on the Entry Clone recombines with attR1 on the Destination Vector, the directionality of the nucleic acid molecule of interest is maintained or may be controlled upon transfer from the Entry Clone into the Destination Vector. Hence, the GATEWAY™ Cloning System provides a powerful and easy method of directional cloning of nucleic acid molecule of interest.
- One-step cloning or subcloning: Mix the Entry Clone and the Destination Vector with Clonase, incubate, and transform.
- Clone PCR products readily by *in vitro* recombination, by adding attB sites to PCR primers. Then directly transfer these Entry Clones into Destination Vectors. This process may also be carried out in one step (see Examples below).
- Powerful selections give high reliability: >90% (and often >99%) of the colonies contain the desired DNA in its new vector.
- One-step conversion of existing standard vectors into GATEWAY™ Cloning System vectors.
- Ideal for large vectors or those with few cloning sites.
- Recombination sites are short (25 bp), and may be engineered to contain no stop codons or secondary structures.
- Reactions may be automated, for high-throughput applications (*e.g.*, for diagnostic purposes or for therapeutic candidate screening).
- The reactions are economical: 0.3 µg of each DNA; no restriction enzymes, phosphatase, ligase, or gel purification. Reactions work well with miniprep DNA.
- Transfer multiple clones, and even libraries, into one or more Destination Vectors, in a single experiment.
- A variety of Destination Vectors may be produced, for applications including, but not limited to:

- Protein expression in *E. coli*: native proteins; fusion proteins with GST, His6, thioredoxin, etc., for purification, or one or more epitope tags; any promoter useful in expressing proteins in *E. coli* may be used, such as ptrc, λP_L , and T7 promoters.
- Protein expression in eukaryotic cells: CMV promoter, baculovirus (with or without His6 tag), Semliki Forest virus, Tet regulation.
- DNA sequencing (all *lac* primers), RNA probes, phagemids (both strands)
- A variety of Entry Vectors (for recombinational cloning entry by standard recombinant DNA methods) may be produced:
 - Strong transcription stop just upstream, for genes toxic to *E. coli*.
 - Three reading frames.
 - With or without TEV protease cleavage site.
 - Motifs for prokaryotic and / or eukaryotic translation.
 - Compatible with commercial cDNA libraries.
- Expression Clone cDNA (attB) libraries, for expression screening, including 2-hybrid libraries and phage display libraries, may also be constructed.

Recombination Site Sequences

In one aspect, the invention relates to nucleic acid molecules, which may or may not be isolated nucleic acid molecules, comprising one or more nucleotide sequences encoding one or more recombination sites or portions thereof. In particular, this aspect of the invention relates to such nucleic acid molecules comprising one or more nucleotide sequences encoding *attB*, *attP*, *attL*, or *attR*, or portions of these recombination site sequences. The invention also relates to mutants, derivatives, and fragments of such nucleic acid molecules. Unless otherwise indicated, all nucleotide sequences that may have been determined by sequencing a DNA molecule herein were determined using manual or automated DNA sequencing, such as dideoxy sequencing, according to methods that are routine to one of ordinary skill in the art (Sanger, F., and Coulson, A.R., *J. Mol. Biol.* 94:444-448 (1975); Sanger, F., *et al.*, *Proc. Natl. Acad. Sci. USA* 74:5463-5467 (1977)). All amino acid sequences of polypeptides encoded by DNA

molecules determined herein were predicted by conceptual translation of a DNA sequence determined as above. Therefore, as is known in the art for any DNA sequence determined by these approaches, any nucleotide sequence determined herein may contain some errors. Nucleotide sequences determined by such methods are typically at least about 90% identical, more typically at least about 95% to at least about 99.9% identical to the actual nucleotide sequence of the sequenced DNA molecule. As is also known in the art, a single insertion or deletion in a determined nucleotide sequence compared to the actual sequence will cause a frame shift in translation of the nucleotide sequence such that the predicted amino acid sequence encoded by a determined nucleotide sequence will be completely different from the amino acid sequence actually encoded by the sequenced DNA molecule, beginning at the point of such an insertion or deletion.

Unless otherwise indicated, each "nucleotide sequence" set forth herein is presented as a sequence of deoxyribonucleotides (abbreviated A, G, C and T). However, by "nucleotide sequence" of a nucleic acid molecule or polynucleotide is intended, for a DNA molecule or polynucleotide, a sequence of deoxyribonucleotides, and for an RNA molecule or polynucleotide, the corresponding sequence of ribonucleotides (A, G, C and U), where each thymidine deoxyribonucleotide (T) in the specified deoxyribonucleotide sequence is replaced by the ribonucleotide uridine (U). Thus, the invention relates to sequences of the invention in the form of DNA or RNA molecules, or hybrid DNA/RNA molecules, and their corresponding complementary DNA, RNA, or DNA/RNA strands.

In a first such aspect, the invention provides nucleic acid molecules comprising one or more nucleotide sequences encoding *attB1*, or mutants, fragments, variants or derivatives thereof. Such nucleic acid molecules may comprise an *attB1* nucleotide sequence having the sequence set forth in Figure 9, such as: ACAAGTTTGTACAAAAAAGCAGGCT, or a nucleotide sequence complementary to the nucleotide sequence set forth in Figure 9 for *attB1*, or mutants, fragments, variants or derivatives thereof. As one of ordinary skill will appreciate, however, certain mutations, insertions, or deletions of one or more bases in the *attB1* sequence contained in the nucleic acid molecules of the invention may be made without compromising the structural and functional

integrity of these molecules; hence, nucleic acid molecules comprising such mutations, insertions, or deletions in the *attB1* sequence are encompassed within the scope of the invention.

In a related aspect, the invention provides nucleic acid molecules comprising one or more nucleotide sequences encoding *attB2*, or mutants, fragments, variants or derivatives thereof. Such nucleic acid molecules may comprise an *attB2* nucleotide sequence having the sequence set forth in Figure 9, such as: ACCCAGCTTTCTTGTACAAAGTGGT, or a nucleotide sequence complementary to the nucleotide sequence set forth in Figure 9 for *attB2*, or mutants, fragments, variants or derivatives thereof. As noted above for *attB1*, certain mutations, insertions, or deletions of one or more bases in the *attB2* sequence contained in the nucleic acid molecules of the invention may be made without compromising the structural and functional integrity of these molecules; hence, nucleic acid molecules comprising such mutations, insertions, or deletions in the *attB2* sequence are encompassed within the scope of the invention.

A recombinant host cell comprising a nucleic acid molecule containing *attB1* and *attB2* sites (the vector pEXP501, also known as pCMVSPORT6; see Figure 48), *E. coli* DB3.1(pCMVSPORT6), was deposited on February 27, 1999, with the Collection, Agricultural Research Culture Collection (NRRL), 1815 North University Street, Peoria, Illinois 61604 USA, as Deposit No. NRRL B-30108. The *attB1* and *attB2* sites within the deposited nucleic acid molecule are contained in nucleic acid cassettes in association with one or more additional functional sequences as described in more detail below.

In another related aspect, the invention provides nucleic acid molecules comprising one or more nucleotide sequences encoding *attP1*, or mutants, fragments, variants or derivatives thereof. Such nucleic acid molecules may comprise an *attP1* nucleotide sequence having the sequence set forth in Figure 9, such as: TACAGGTCACCTAATACCATCTAAGTAGTTGATTCATAGTGA-CTGGATATGTTGTGTTTTACAGTATTATGTAGTCTGTTTTTTAT-GCAAAATCTAATTTAATATATTGATATTTATATCATTTTACGTT-TCTCGTTCAGCTTTTTTTGTACAAAGTTGGCATTATAAAAAAGCATTG-CTCATCAATTTGTTGCAACGAACAGGTCACCTATCAGTCAAAATAA-

AATCATTATTTG, or a nucleotide sequence complementary to the nucleotide sequence set forth in Figure 9 for *attP1*, or mutants, fragments, variants or derivatives thereof. As noted above for *attB1*, certain mutations, insertions, or deletions of one or more bases in the *attP1* sequence contained in the nucleic acid molecules of the invention may be made without compromising the structural and functional integrity of these molecules; hence, nucleic acid molecules comprising such mutations, insertions, or deletions in the *attP1* sequence are encompassed within the scope of the invention.

In another related aspect, the invention provides nucleic acid molecules comprising one or more nucleotide sequences encoding *attP2*, or mutants, fragments, variants or derivatives thereof. Such nucleic acid molecules may comprise an *attP2* nucleotide sequence having the sequence set forth in Figure 9, such as: CAAATAATGATTTTATTTTGACTGATAGTGACCTGTTTCGTTG-CAACAAATTGATAAGCAATGCTTTCTTATAATGCCAACTTT-GTACAAGAAAGCTGAACGAGAAACGTAAAATGATA-TAAATATCAATATATTAAATTAGATTTTGCATAAAAAACAG-ACTACATAATACTGTAAAACACAACATATCCAGTCACTATGAATCAACTACTTAGATGGTATTAGTGACCTGTA, or a nucleotide sequence complementary to the nucleotide sequence set forth in Figure 9 for *attP2*, or mutants, fragments, variants or derivatives thereof. As noted above for *attB1*, certain mutations, insertions, or deletions of one or more bases in the *attP2* sequence contained in the nucleic acid molecules of the invention may be made without compromising the structural and functional integrity of these molecules; hence, nucleic acid molecules comprising such mutations, insertions, or deletions in the *attP2* sequence are encompassed within the scope of the invention.

A recombinant host cell comprising a nucleic acid molecule (the *attP* vector pDONR201, also known as pENTR21-*attPkan* or pAttPkan; see Figure 49) containing *attP1* and *attP2* sites, *E. coli* DB3.1(pAttPkan) (also called *E. coli* DB3.1(pAHKan)), was deposited on February 27, 1999, with the Collection, Agricultural Research Culture Collection (NRRL), 1815 North University Street, Peoria, Illinois 61604 USA, as Deposit No. NRRL B-30099. The *attP1* and *attP2* sites within the deposited nucleic acid molecule are contained in nucleic acid

cassettes in association with one or more additional functional sequences as described in more detail below.

In another related aspect, the invention provides nucleic acid molecules comprising one or more nucleotide sequences encoding *attR1*, or mutants, fragments, variants or derivatives thereof. Such nucleic acid molecules may comprise an *attR1* nucleotide sequence having the sequence set forth in Figure 9, such as: ACAAGTTTGTACAAAAAGCTGAACGAG-AAACGTAAAATGATATAAATATCAATATATTAAATTAGATTTTGCAT-AAAAACAGACTACATAATACTGTAAACACAACATATCCAGTCA-CTATG, or a nucleotide sequence complementary to the nucleotide sequence set forth in Figure 9 for *attR1*, or mutants, fragments, variants or derivatives thereof. As noted above for *attB1*, certain mutations, insertions, or deletions of one or more bases in the *attR1* sequence contained in the nucleic acid molecules of the invention may be made without compromising the structural and functional integrity of these molecules; hence, nucleic acid molecules comprising such mutations, insertions, or deletions in the *attR1* sequence are encompassed within the scope of the invention.

In another related aspect, the invention provides nucleic acid molecules comprising one or more nucleotide sequences encoding *attR2*, or mutants, fragments, variants or derivatives thereof. Such nucleic acid molecules may comprise an *attR2* nucleotide sequence having the sequence set forth in Figure 9, such as: GCAGGTCGACCATAGTGACTGGATAT-GTTGTGTTTTACAGTATTATGTAGTCTGTTTTTATGCAAATCTA-ATTTAATATATTGATATTTATATCATTTTACGTTTCTCGTTCAGCTT-TCTTGTACAAAGTGGT, or a nucleotide sequence complementary to the nucleotide sequence set forth in Figure 9 for *attR2*, or mutants, fragments, variants or derivatives thereof. As noted above for *attB1*, certain mutations, insertions, or deletions of one or more bases in the *attR2* sequence contained in the nucleic acid molecules of the invention may be made without compromising the structural and functional integrity of these molecules; hence, nucleic acid molecules comprising such mutations, insertions, or deletions in the *attR2* sequence are encompassed within the scope of the invention.

Recombinant host cell strains containing attR1 sites apposed to cloning sites in reading frame A, reading frame B, and reading frame C, *E. coli* DB3.1(pEZC15101) (reading frame A; see Figure 64A), *E. coli* DB3.1(pEZC15102) (reading frame B; see Figure 64B), and *E. coli* DB3.1(pEZC15103) (reading frame C; see Figure 64C), and containing corresponding attR2 sites, were deposited on February 27, 1999, with the Collection, Agricultural Research Culture Collection (NRRL), 1815 North University Street, Peoria, Illinois 61604 USA, as Deposit Nos. NRRL B-30103, NRRL B-30104, and NRRL B-30105, respectively. The attR1 and attR2 sites within the deposited nucleic acid molecules are contained in nucleic acid cassettes in association with one or more additional functional sequences as described in more detail below.

In another related aspect, the invention provides nucleic acid molecules comprising one or more nucleotide sequences encoding attL1, or mutants, fragments, variants and derivatives thereof. Such nucleic acid molecules may comprise an attL1 nucleotide sequence having the sequence set forth in Figure 9, such as: CAA ATA ATG ATT TTA TTT TGA CTG ATA GTG ACC TGT TCG TTG CAA CAA ATT GAT AAG CAA TGC TTT TTT ATA ATG CCA ACT TTG TAC AAA AAA GCA GGC T, or a nucleotide sequence complementary to the nucleotide sequence set forth in Figure 9 for attL1, or mutants, fragments, variants or derivatives thereof. As noted above for attB1, certain mutations, insertions, or deletions of one or more bases in the attL1 sequence contained in the nucleic acid molecules of the invention may be made without compromising the structural and functional integrity of these molecules; hence, nucleic acid molecules comprising such mutations, insertions, or deletions in the attL1 sequence are encompassed within the scope of the invention.

In another related aspect, the invention provides nucleic acid molecules comprising one or more nucleotide sequences encoding attL2, or mutants, fragments, variants and derivatives thereof. Such nucleic acid molecules may comprise an attL2 nucleotide sequence having the sequence set forth in Figure 9, such as: C AAA TAA TGA TTT TAT TTT GAC TGA TAG TGA CCT GTT CGT TGC AAC AAA TTG ATA AGC AAT GCT TTC TTA TAA TGC CAA

CTT TGT ACA AGA AAG CTG GGT, or a nucleotide sequence complementary to the nucleotide sequence set forth in Figure 9 for *attL2*, or mutants, fragments, variants or derivatives thereof. As noted above for *attB1*, certain mutations, insertions, or deletions of one or more bases in the *attL2* sequence contained in the nucleic acid molecules of the invention may be made without compromising the structural and functional integrity of these molecules; hence, nucleic acid molecules comprising such mutations, insertions, or deletions in the *attL2* sequence are encompassed within the scope of the invention.

Recombinant host cell strains containing *attL1* sites apposed to cloning sites in reading frame A, reading frame B, and reading frame C, *E. coli* DB3.1(pENTR1A) (reading frame A; see Figure 10), *E. coli* DB3.1(pENTR2B) (reading frame B; see Figure 11), and *E. coli* DB3.1(pENTR3C) (reading frame C; see Figure 12), and containing corresponding *attL2* sites, were deposited on February 27, 1999, with the Collection, Agricultural Research Culture Collection (NRRL), 1815 North University Street, Peoria, Illinois 61604 USA, as Deposit Nos. NRRL B-30100, NRRL B-30101, and NRRL B-30102, respectively. The *attL1* and *attL2* sites within the deposited nucleic acid molecules are contained in nucleic acid cassettes in association with one or more additional functional sequences as described in more detail below.

Each of the recombination site sequences described herein or portions thereof, or the nucleotide sequence cassettes contained in the deposited clones, may be cloned or inserted into a vector of interest (for example, using the recombinational cloning methods described herein and/or standard restriction cloning techniques that are routine in the art) to generate, for example, Entry Vectors or Destination Vectors which may be used to transfer a desired segment of a nucleic acid molecule of interest (*e.g.*, a gene, cDNA molecule, or cDNA library) into a desired vector or into a host cell.

Using the information provided herein, such as the nucleotide sequences for the recombination site sequences described herein, an isolated nucleic acid molecule of the present invention encoding one or more recombination sites or portions thereof may be obtained using standard cloning and screening procedures, such as those for cloning cDNAs using mRNA as starting material. Preferred such

methods include PCR-based cloning methods, such as reverse transcriptase-PCR (RT-PCR) using primers such as those described herein and in the Examples below. Alternatively, vectors comprising the cassettes containing the recombination site sequences described herein are available commercially from Life Technologies, Inc. (Rockville, MD).

The invention is also directed to nucleic acid molecules comprising one or more of the recombination site sequences or portions thereof and one or more additional nucleotide sequences, which may encode functional or structural sites such as one or more multiple cloning sites, one or more transcription termination sites, one or more transcriptional regulatory sequences (which may be promoters, enhancers, repressors, and the like), one or more translational signals (e.g., secretion signal sequences), one or more origins of replication, one or more fusion partner peptides (particularly glutathione S-transferase (GST), hexahistidine (His₆), and thioredoxin (Trx)), one or more selection markers or modules, one or more nucleotide sequences encoding localization signals such as nuclear localization signals or secretion signals, one or more origins of replication, one or more protease cleavage sites, one or more genes or portions of genes encoding a protein or polypeptide of interest, and one or more 5' polynucleotide extensions (particularly an extension of guanine residues ranging in length from about 1 to about 20, from about 2 to about 15, from about 3 to about 10, from about 4 to about 10, and most preferably an extension of 4 or 5 guanine residues at the 5' end of the recombination site nucleotide sequence. The one or more additional functional or structural sequences may or may not flank one or more of the recombination site sequences contained on the nucleic acid molecules of the invention.

In some nucleic acid molecules of the invention, the one or more nucleotide sequences encoding one or more additional functional or structural sites may be operably linked to the nucleotide sequence encoding the recombination site. For example, certain nucleic acid molecules of the invention may have a promoter sequence operably linked to a nucleotide sequence encoding a recombination site or portion thereof of the invention, such as a T7 promoter, a phage lambda PL

promoter, an *E. coli lac*, *trp* or *tac* promoter, and other suitable promoters which will be familiar to the skilled artisan.

Nucleic acid molecules of the present invention, which may be isolated nucleic acid molecules, may be in the form of RNA, such as mRNA, or in the form of DNA, including, for instance, cDNA and genomic DNA obtained by cloning or produced synthetically, or in the form of DNA-RNA hybrids. The nucleic acid molecules of the invention may be double-stranded or single-stranded. Single-stranded DNA or RNA may be the coding strand, also known as the sense strand, or it may be the non-coding strand, also referred to as the anti-sense strand. The nucleic acid molecules of the invention may also have a number of topologies, including linear, circular, coiled, or supercoiled.

By "isolated" nucleic acid molecule(s) is intended a nucleic acid molecule, DNA or RNA, which has been removed from its native environment. For example, recombinant DNA molecules contained in a vector are considered isolated for the purposes of the present invention. Further examples of isolated DNA molecules include recombinant DNA molecules maintained in heterologous host cells, and those DNA molecules purified (partially or substantially) from a solution whether produced by recombinant DNA or synthetic chemistry techniques. Isolated RNA molecules include *in vivo* or *in vitro* RNA transcripts of the DNA molecules of the present invention.

The present invention further relates to mutants, fragments, variants and derivatives of the nucleic acid molecules of the present invention, which encode portions, analogs or derivatives of one or more recombination sites. Variants may occur naturally, such as a natural allelic variant. By an "allelic variant" is intended one of several alternate forms of a gene occupying a given locus on a chromosome of an organism (*see* Lewin, B., ed., *Genes II*, , John Wiley & Sons, New York (1985)). Non-naturally occurring variants may be produced using art-known mutagenesis techniques, such as those described hereinbelow.

Such variants include those produced by nucleotide substitutions, deletions or additions or portions thereof, or combinations thereof. The substitutions, deletions or additions may involve one or more nucleotides. The variants may be altered in coding regions, non-coding regions, or both. Alterations in the coding

regions may produce conservative or non-conservative amino acid substitutions, deletions or additions. Especially preferred among these are silent substitutions, additions and deletions, which do not alter the properties and activities of the encoded polypeptide(s) or portions thereof, and which also do not substantially alter the reactivities of the recombination site nucleic acid sequences in recombination reactions. Also especially preferred in this regard are conservative substitutions.

Particularly preferred mutants, fragments, variants, and derivatives of the nucleic acid molecules of the invention include, but are not limited to, insertions, deletions or substitutions of one or more nucleotide bases within the 15 bp core region (GCTTTTTTATACTAA) which is identical in all four wildtype lambda *att* sites, *attB*, *attP*, *attL* and *attR* (see U.S. Application Nos. 08/663,002, filed June 7, 1996 (now U.S. Patent No. 5,888,732), 09/005,476, filed January 12, 1998, and 09/177,387, filed October 23, 1998, which describes the core region in further detail, and the disclosures of which are incorporated herein by reference in their entireties). Analogously, the core regions in *attB1*, *attP1*, *attL1* and *attR1* are identical to one another, as are the core regions in *attB2*, *attP2*, *attL2* and *attR2*. Particularly preferred in this regard are nucleic acid molecules comprising insertions, deletions or substitutions of one or more nucleotides within the seven bp overlap region (TTTATAC, which is defined by the cut sites for the integrase protein and is the region where strand exchange takes place) that occurs within this 15 bp core region (GCTTTTTTTTTATACTAA). Examples of such preferred mutants, fragments, variants and derivatives according to this aspect of the invention include, but are not limited to, nucleic acid molecules in which the thymine at position 1 of the seven bp overlap region has been deleted or substituted with a guanine, cytosine, or adenine; in which the thymine at position 2 of the seven bp overlap region has been deleted or substituted with a guanine, cytosine, or adenine; in which the thymine at position 3 of the seven bp overlap region has been deleted or substituted with a guanine, cytosine, or adenine; in which the adenine at position 4 of the seven bp overlap region has been deleted or substituted with a guanine, cytosine, or thymine; in which the thymine at position 5 of the seven bp overlap region has been deleted or substituted with a

5 guanine, cytosine, or adenine; in which the adenine at position 6 of the seven bp overlap region has been deleted or substituted with a guanine, cytosine, or thymine; and in which the cytosine at position 7 of the seven bp overlap region has been deleted or substituted with a guanine, thymine, or adenine; or any combination of one or more such deletions and/or substitutions within this seven bp overlap region. As described in detail in Example 21 herein, mutants of the nucleic acid molecules of the invention in which substitutions have been made within the first three positions of the seven bp overlap (TTTAATAC) have been found in the present invention to strongly affect the specificity of recombination, mutant nucleic acid molecules in which substitutions have been made in the last 10 four positions (TTTAATAC) only partially alter recombination specificity, and mutant nucleic acid molecules comprising nucleotide substitutions outside of the seven bp overlap, but elsewhere within the 15 bp core region, do not affect specificity of recombination but do influence the efficiency of recombination.

15 Hence, in an additional aspect, the present invention is also directed to nucleic acid molecules comprising one or more recombination site nucleotide sequences that affect recombination specificity, particularly one or more nucleotide sequences that may correspond substantially to the seven base pair overlap within the 15 bp core region, having one or more mutations that affect recombination specificity. Particularly preferred such molecules may comprise a consensus 20 sequence (described in detail in Example 21 herein) such as NNNATAC, wherein "N" refers to any nucleotide (*i.e.*, may be A, G, T/U or C), with the proviso that if one of the first three nucleotides in the consensus sequence is a T/U, then at least one of the other two of the first three nucleotides is not a T/U.

25 In a related aspect, the present invention is also directed to nucleic acid molecules comprising one or more recombination site nucleotide sequences that enhance recombination efficiency, particularly one or more nucleotide sequences that may correspond substantially to the core region and having one or more mutations that enhance recombination efficiency. By sequences or mutations that 30 "enhance recombination efficiency" is meant a sequence or mutation in a recombination site, preferably in the core region (*e.g.*, the 15 bp core region of *att* recombination sites), that results in an increase in cloning efficiency (typically

measured by determining successful cloning of a test sequence, *e.g.*, by determining CFU/ml for a given cloning mixture) when recombining molecules comprising the mutated sequence or core region as compared to molecules that do not comprise the mutated sequence or core region (*e.g.*, those comprising a wildtype recombination site core region sequence). More specifically, whether or not a given sequence or mutation enhances recombination efficiency may be determined using the sequence or mutation in recombinational cloning as described herein, and determining whether the sequence or mutation provides enhanced recombinational cloning efficiency when compared to a non-mutated (*e.g.*, wildtype) sequence. Methods of determining preferred cloning efficiency-enhancing mutations for a number of recombination sites, particularly for *att* recombination sites, are described herein, for example in Examples 22-25. Examples of preferred such mutant recombination sites include but are not limited to the *attL* consensus core sequence of caactnntnnnannaagttg (wherein "n" represents any nucleotide), for example the *attL5* sequence agcctgctttattataactaagttggcatta and the *attL6* sequence agcctgcttttttatattaagttggcatta; the *attB1.6* sequence ggggacaactttgtacaaaaagttggct; the *attB2.2* sequence ggggacaactttgtacaagaaagctgggt; and the *attB2.10* sequence ggggacaactttgtacaagaaagttgggt. Those of skill in the art will appreciate that, in addition to the core region, other portions of the *att* site may affect the efficiency of recombination. There are five so-called arm binding sites for the integrase protein in the bacteriophage lambda *attP* site, two in *attR* (P1 and P2), and three in *attL* (P'1, P'2 and P'3). Compared to the core binding sites, the integrase protein binds to arm sites with high affinity and interacts with core and arm sites through two different domains of the protein. As with the core binding site a consensus sequence for the arm binding site consisting of C/AAGTCACTAT has been inferred from sequence comparison of the five arm binding sites and seven non-*att* sites (Ross and Landy, *Proc. Natl. Acad. Sci. USA* 79:7724-7728 (1982)). Each arm site has been mutated and tested for its effect in the excision and integration reactions (Numrych *et al.*, *Nucl. Acids Res.* 18:3953 (1990)). Hence, specific sites are utilized in each reaction in different ways, namely, the P1 and P'3

sites are essential for the integration reaction whereas the other three sites are dispensable to the integration reaction to varying degrees. Similarly, the P2, P'1 and P'2 sites are most important for the excision reaction, whereas P1 and P'3 are completely dispensable. Interestingly, when P2 is mutated the integration reaction occurs more efficiently than with the wild type attP site. Similarly, when P1 and P'3 are mutated the excision reaction occurs more efficiently. The stimulatory effect of mutating integrase arm binding sites can be explained by removing sites that compete or inhibit a specific recombination pathway or that function in a reaction that converts products back to starting substrates. In fact there is evidence for an XIS-independent LR reaction (Abremski and Gottesman, *J. Mol. Biol.* 153:67-78 (1981)). Thus, in addition to modifications in the core region of the att site, the present invention contemplates the use of att sites containing one or more modifications in the integrase arm-type binding sites. In some preferred embodiments, one or more mutations may be introduced into one or more of the P1, P'1, P2, P'2 and P'3 sites. In some preferred embodiments, multiple mutations may be introduced into one or more of these sites. Preferred such mutations include those which increase the recombination *in vitro*. For example, in some embodiments mutations may be introduced into the arm-type binding sites such that integrative recombination, corresponding to the BP reaction, is enhanced. In other embodiments, mutations may be introduced into the arm-type binding sites such that excisive recombination, corresponding to the LR reaction, is enhanced. Of course, based on the guidance contained herein, particularly in the construction and evaluation of effects of mutated recombination sites upon recombinational specificity and efficiency, analogous mutated or engineered sequences may be produced for other recombination sites described herein (including but not limited to *lox*, FRT, and the like) and used in accordance with the invention. For example, much like the mutagenesis strategy used to select core binding sites that enhance recombination efficiency, similar strategies can be employed to select changes in the arms of attP, attL and attR, and in analogous sequences in other recombination sites such as *lox*, FRT and the like, that enhance recombination efficiency. Hence, the construction and evaluation of such mutants is well within the abilities of those of ordinary skill in the art without undue experimentation.

One suitable methodology for preparing and evaluating such mutations is found in Numrych, *et al.*, (1990) *Nucleic Acids Research* 18(13): 3953-3959.

Other mutant sequences and nucleic acid molecules that may be suitable to enhance recombination efficiency will be apparent from the description herein, or may be easily determined by one of ordinary skill using only routine experimentation in molecular biology in view of the description herein and information that is readily available in the art

Since the genetic code is well known in the art, it is also routine for one of ordinary skill in the art to produce degenerate variants of the nucleic acid molecules described herein without undue experimentation. Hence, nucleic acid molecules comprising degenerate variants of nucleic acid sequences encoding the recombination sites described herein are also encompassed within the scope of the invention.

Further embodiments of the invention include isolated nucleic acid molecules comprising a polynucleotide having a nucleotide sequence at least 50% identical, at least 60% identical, at least 70% identical, at least 75% identical, at least 80% identical, at least 85% identical, at least 90% identical, and more preferably at least 95%, 96%, 97%, 98% or 99% identical to the nucleotide sequences of the seven bp overlap region within the 15 bp core region of the recombination sites described herein, or the nucleotide sequences of *attB1*, *attB2*, *attP1*, *attP2*, *attL1*, *attL2*, *attR1* or *attR2* as set forth in Figure 9 (or portions thereof), or a nucleotide sequence complementary to any of these nucleotide sequences, or fragments, variants, mutants, and derivatives thereof.

By a polynucleotide having a nucleotide sequence at least, for example, 95% "identical" to a reference nucleotide sequence encoding a particular recombination site or portion thereof is intended that the nucleotide sequence of the polynucleotide is identical to the reference sequence except that the polynucleotide sequence may include up to five point mutations (e.g., insertions, substitutions, or deletions) per each 100 nucleotides of the reference nucleotide sequence encoding the recombination site. For example, to obtain a polynucleotide having a nucleotide sequence at least 95% identical to a reference *attB1* nucleotide sequence, up to 5% of the nucleotides in the *attB1* reference sequence may be

deleted or substituted with another nucleotide, or a number of nucleotides up to 5% of the total nucleotides in the *attB1* reference sequence may be inserted into the *attB1* reference sequence. These mutations of the reference sequence may occur at the 5' or 3' terminal positions of the reference nucleotide sequence or
5 anywhere between those terminal positions, interspersed either individually among nucleotides in the reference sequence or in one or more contiguous groups within the reference sequence.

As a practical matter, whether any particular nucleic acid molecule is at least 50%, 60%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98% or 99% identical
10 to, for instance, a given recombination site nucleotide sequence or portion thereof can be determined conventionally using known computer programs such as DNAsis software (Hitachi Software, San Bruno, California) for initial sequence alignment followed by ESEE version 3.0 DNA/protein sequence software (cabot@trog.mbb.sfu.ca) for multiple sequence alignments. Alternatively, such
15 determinations may be accomplished using the BESTFIT program (Wisconsin Sequence Analysis Package, Genetics Computer Group, University Research Park, 575 Science Drive, Madison, WI 53711), which employs a local homology algorithm (Smith and Waterman, *Advances in Applied Mathematics* 2: 482-489 (1981)) to find the best segment of homology between two sequences. When
20 using DNAsis, ESEE, BESTFIT or any other sequence alignment program to determine whether a particular sequence is, for instance, 95% identical to a reference sequence according to the present invention, the parameters are set such that the percentage of identity is calculated over the full length of the reference nucleotide sequence and that gaps in homology of up to 5% of the total number
25 of nucleotides in the reference sequence are allowed.

The present invention is directed to nucleic acid molecules at least 50%, 60%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98% or 99% identical to the *attB1*, *attB2*, *attP1*, *attP2*, *attL1*, *attL2*, *attR1* or *attR2* nucleotide sequences as set forth in Figure 9, or to the nucleotide sequence of the deposited clones, irrespective of
30 whether they encode particular functional polypeptides. This is because even where a particular nucleic acid molecule does not encode a particular functional polypeptide, one of skill in the art would still know how to use the nucleic acid

molecule, for instance, as a hybridization probe or a polymerase chain reaction (PCR) primer.

5 Mutations can also be introduced into the recombination site nucleotide sequences for enhancing site specific recombination or altering the specificities of the reactants, etc. Such mutations include, but are not limited to: recombination sites without translation stop codons that allow fusion proteins to be encoded; recombination sites recognized by the same proteins but differing in base sequence such that they react largely or exclusively with their homologous partners allowing multiple reactions to be contemplated; and mutations that prevent hairpin
10 formation of recombination sites. Which particular reactions take place can be specified by which particular partners are present in the reaction mixture.

There are well known procedures for introducing specific mutations into nucleic acid sequences. A number of these are described in Ausubel, F.M. *et al.*, *Current Protocols in Molecular Biology*, Wiley Interscience, New York (1989-
15 1996). Mutations can be designed into oligonucleotides, which can be used to modify existing cloned sequences, or in amplification reactions. Random mutagenesis can also be employed if appropriate selection methods are available to isolate the desired mutant DNA or RNA. The presence of the desired mutations can be confirmed by sequencing the nucleic acid by well known
20 methods.

The following non-limiting methods can be used to modify or mutate a given nucleic acid molecule encoding a particular recombination site to provide mutated sites that can be used in the present invention:

- 25 1. By recombination of two parental DNA sequences by site-specific (e.g. attL and attR to give attP) or other (e.g. homologous) recombination mechanisms where the parental DNA segments contain one or more base alterations resulting in the final mutated nucleic acid molecule;
2. By mutation or mutagenesis (site-specific, PCR, random, spontaneous, etc) directly of the desired nucleic acid molecule;
- 30 3. By mutagenesis (site-specific, PCR, random, spontaneous, etc) of parental DNA sequences, which are recombined to generate a desired nucleic acid molecule;

4. By reverse transcription of an RNA encoding the desired core sequence;
and
5. By *de novo* synthesis (chemical synthesis) of a sequence having the desired
base changes, or random base changes followed by sequencing or
functional analysis according to methods that are routine in the art.

The functionality of the mutant recombination sites can be demonstrated in ways that depend on the particular characteristic that is desired. For example, the lack of translation stop codons in a recombination site can be demonstrated by expressing the appropriate fusion proteins. Specificity of recombination between homologous partners can be demonstrated by introducing the appropriate molecules into *in vitro* reactions, and assaying for recombination products as described herein or known in the art. Other desired mutations in recombination sites might include the presence or absence of restriction sites, translation or transcription start signals, protein binding sites, particular coding sequences, and other known functionalities of nucleic acid base sequences. Genetic selection schemes for particular functional attributes in the recombination sites can be used according to known method steps. For example, the modification of sites to provide (from a pair of sites that do not interact) partners that do interact could be achieved by requiring deletion, via recombination between the sites, of a DNA sequence encoding a toxic substance. Similarly, selection for sites that remove translation stop sequences, the presence or absence of protein binding sites, etc., can be easily devised by those skilled in the art.

Accordingly, the present invention also provides a nucleic acid molecule, comprising at least one DNA segment having at least one, and preferably at least two, engineered recombination site nucleotide sequences of the invention flanking a selectable marker and/or a desired DNA segment, wherein at least one of said recombination site nucleotide sequences has at least one engineered mutation that enhances recombination *in vitro* in the formation of a Cointegrate DNA or a Product DNA. Such engineered mutations may be in the core sequence of the recombination site nucleotide sequence of the invention; see U.S. Application Nos. 08/486,139, filed June 7, 1995, 08/663,002, filed June 7, 1996 (now U.S. Patent No. 5,888,732), 09/005,476, filed January 12, 1998, and 09/177,387, filed

October 23, 1998, the disclosures of which are all incorporated herein by reference in their entireties.

While in the preferred embodiment the recombination sites differ in sequence and do not interact with each other, it is recognized that sites comprising the same sequence, which may interact with each other, can be manipulated or engineered to inhibit recombination with each other. Such conceptions are considered and incorporated herein. For example, a protein binding site (*e.g.*, an antibody-binding site, a histone-binding site, an enzyme-binding site, or a binding site for any nucleic acid molecule-binding protein) can be engineered adjacent to one of the sites. In the presence of the protein that recognizes the engineered site, the recombinase fails to access the site and another recombination site in the nucleic acid molecule is therefore used preferentially. In the cointegrate this site can no longer react since it has been changed, *e.g.*, from attB to attL. During or upon resolution of the cointegrate, the protein can be inactivated (*e.g.*, by antibody, heat or a change of buffer) and the second site can undergo recombination.

The nucleic acid molecules of the invention can have at least one mutation that confers at least one enhancement of said recombination, said enhancement selected from the group consisting of substantially (i) favoring integration; (ii) favoring recombination; (iii) relieving the requirement for host factors; (iv) increasing the efficiency of said Cointegrate DNA or Product DNA formation; (v) increasing the specificity of said Cointegrate DNA or Product DNA formation; and (vi) adding or deleting protein binding sites.

In other embodiments, the nucleic acid molecules of the invention may be PCR primer molecules, which comprise one or more of the recombination site sequences described herein or portions thereof, particularly those shown in Figure 9 (or sequences complementary to those shown in Figure 9), or mutants, fragments, variants or derivatives thereof, attached at the 3' end to a target-specific template sequence which specifically interacts with a target nucleic acid molecule which is to be amplified. Primer molecules according to this aspect of the invention may further comprise one or more, (*e.g.*, 1, 2, 3, 4, 5, 10, 20, 25, 50, 100, 500, 1000, or more) additional bases at their 5' ends, and preferably comprise one or more (particularly four or five) additional bases, which are preferably

guanines, at their 5' ends, to increase the efficiency of the amplification products incorporating the primer molecules in the recombinational cloning system of the invention. Such nucleic acid molecules and primers are described in detail in the examples herein, particularly in Examples 22-25.

5 Certain primers of the invention may comprise one or more nucleotide deletions in the *attB1*, *attB2*, *attP1*, *attP2*, *attL1*, *attL2*, *attR1* or *attR2* sequences as set forth in Figure 9. In one such aspect, for example, *attB2* primers may be constructed in which one or more of the first four nucleotides at the 5' end of the *attB2* sequence shown in Figure 9 have been deleted. Primers according to this aspect of the invention may therefore have the sequence:

10 (*attB2*(-1)): CCCAGCTTTCTTGTACAAAGTGGTnnnnnnnnnnnnnn . . . n

(*attB2*(-2)): CCAGCTTTCTTGTACAAAGTGGTnnnnnnnnnnnnnn . . . n

(*attB2*(-3)): CAGCTTTCTTGTACAAAGTGGTnnnnnnnnnnnnnn . . . n

(*attB2*(-4)): AGCTTTCTTGTACAAAGTGGTnnnnnnnnnnnnnn . . . n,

15 wherein "nnnnnnnnnnnn . . . n" at the 3' end of the primer represents a target-specific sequence of any length, for example from one base up to all of the bases of a target nucleic acid molecule (*e.g.*, a gene) or a portion thereof, the sequence and length which will depend upon the identity of the target nucleic acid molecule which is to be amplified.

20 The primer nucleic acid molecules according to this aspect of the invention may be produced synthetically by attaching the recombination site sequences depicted in Figure 9, or portions thereof, to the 5' end of a standard PCR target-specific primer according to methods that are well-known in the art. Alternatively, additional primer nucleic acid molecules of the invention may be produced synthetically by adding one or more nucleotide bases, which preferably correspond to one or more, preferably five or more, and more preferably six or more, contiguous nucleotides of the *att* nucleotide sequences described herein (*see, e.g.*, Example 20 herein; *see also* U.S. Application Nos. 08/663,002, filed June 7, 1996 (now U.S. Patent No. 5,888,732), 09/005,476, filed January 12, 1998, and 09/177,387, filed October 23, 1998, the disclosures of which are all incorporated herein by reference in their entireties), to the 5' end of a standard PCR target-specific primer according to methods that are well-known in the art, to provide

25

30

primers having the specific nucleotide sequences described herein. As noted above, primer nucleic acid molecules according to this aspect of the invention may also optionally comprise one, two, three, four, five, or more additional nucleotide bases at their 5' ends, and preferably will comprise four or five guanines at their 5' ends. In one particularly preferred such aspect, the primer nucleic acid molecules of the invention may comprise one or more, preferably five or more, more preferably six or more, still more preferably 6-18 or 6-25, and most preferably 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24 or 25, contiguous nucleotides or bp of the *attB1* or *attB2* nucleotide sequences depicted in Figure 9 (or nucleotides complementary thereto), linked to the 5' end of a target-specific (*e.g.*, a gene-specific) primer molecule. Primer nucleic acid molecules according to this aspect of the invention include, but are not limited to, *attB1*- and *attB2*-derived primer nucleic acid molecules having the following nucleotide sequences:

ACAAGTTTGTACAAAAAAGCAGGCT-nnnnnnnnnnnnn . . . n
ACCACTTTGTACAAGAAAGCTGGGT-nnnnnnnnnnnnn . . . n
TGTACAAAAAAGCAGGCT-nnnnnnnnnnnnn . . . n
TGTACAAGAAAGCTGGGT-nnnnnnnnnnnnn . . . n
ACAAAAAAGCAGGCT-nnnnnnnnnnnnn . . . n
ACAAGAAAGCTGGGT-nnnnnnnnnnnnn . . . n
AAAAAGCAGGCT-nnnnnnnnnnnnn . . . n
AGAAAGCTGGGT-nnnnnnnnnnnnn . . . n
AAAAGCAGGCT-nnnnnnnnnnnnn . . . n
GAAAGCTGGGT-nnnnnnnnnnnnn . . . n
AAAGCAGGCT-nnnnnnnnnnnnn . . . n
AAAGCTGGGT-nnnnnnnnnnnnn . . . n
AAGCAGGCT-nnnnnnnnnnnnn . . . n
AAGCTGGGT-nnnnnnnnnnnnn . . . n
AGCAGGCT-nnnnnnnnnnnnn . . . n
AGCTGGGT-nnnnnnnnnnnnn . . . n
GCAGGCT-nnnnnnnnnnnnn . . . n
GCTGGGT-nnnnnnnnnnnnn . . . n

CAGGCT-nnnnnnnnnnnnn . . . n

CTGGGT-nnnnnnnnnnnnn . . . n,

wherein "nnnnnnnnnnnn . . . n" at the 3' end of the primer represents a target-specific sequence of any length, for example from one base up to all of the bases of a target nucleic acid molecule (e.g., a gene) or a portion thereof, the sequence and length which will depend upon the identity of the target nucleic acid molecule which is to be amplified.

Of course, it will be apparent to one of ordinary skill from the teachings contained herein that additional primer nucleic acid molecules analogous to those specifically described herein may be produced using one or more, preferably five or more, more preferably six or more, still more preferably ten or more, 15 or more, 20 or more, 25 or more, 30 or more, etc. (through to and including all) of the contiguous nucleotides or bp of the *attP1*, *attP2*, *attL1*, *attL2*, *attR1* or *attR2* nucleotide sequences depicted in Figure 9 (or nucleotides complementary thereto), linked to the 5' end of a target-specific (e.g., a gene-specific) primer molecule. As noted above, such primer nucleic acid molecules may optionally further comprise one, two, three, four, five, or more additional nucleotide bases at their 5' ends, and preferably will comprise four guanines at their 5' ends. Other primer molecules comprising the *attB1*, *attB2*, *attP1*, *attP2*, *attL1*, *attL2*, *attR1* and *attR2* sequences depicted in Figure 9, or portions thereof, may be made by one of ordinary skill without resorting to undue experimentation in accordance with the guidance provided herein.

The primers of the invention described herein are useful in producing PCR fragments having a nucleic acid molecule of interest flanked at each end by a recombination site sequence (as described in detail below in Example 9), for use in cloning of PCR-amplified DNA fragments using the recombination system of the invention (as described in detail below in Examples 8, 19 and 21-25).

Vectors

The invention also relates to vectors comprising one or more of the nucleic acid molecules of the invention, as described herein. In accordance with the invention, any vector may be used to construct the vectors of the invention. In

particular, vectors known in the art and those commercially available (and variants or derivatives thereof) may in accordance with the invention be engineered to include one or more nucleic acid molecules encoding one or more recombination sites (or portions thereof), or mutants, fragments, or derivatives thereof, for use in the methods of the invention. Such vectors may be obtained from, for example, Vector Laboratories Inc., InVitrogen, Promega, Novagen, New England Biolabs, Clontech, Roche, Pharmacia, EpiCenter, OriGenes Technologies Inc., Stratagene, Perkin Elmer, Pharmingen, Life Technologies, Inc., and Research Genetics. Such vectors may then for example be used for cloning or subcloning nucleic acid molecules of interest. General classes of vectors of particular interest include prokaryotic and/or eukaryotic cloning vectors, Expression Vectors, fusion vectors, two-hybrid or reverse two-hybrid vectors, shuttle vectors for use in different hosts, mutagenesis vectors, transcription vectors, vectors for receiving large inserts and the like.

Other vectors of interest include viral origin vectors (M13 vectors, bacterial phage λ vectors, bacteriophage P1 vectors, adenovirus vectors, herpesvirus vectors, retrovirus vectors, phage display vectors, combinatorial library vectors), high, low, and adjustable copy number vectors, vectors which have compatible replicons for use in combination in a single host (pACYC184 and pBR322) and eukaryotic episomal replication vectors (pCDM8).

Particular vectors of interest include prokaryotic Expression Vectors such as pcDNA II, pSL301, pSE280, pSE380, pSE420, pTrcHisA, B, and C, pRSET A, B, and C (Invitrogen, Inc.), pGEMEX-1, and pGEMEX-2 (Promega, Inc.), the pET vectors (Novagen, Inc.), pTrc99A, pKK223-3, the pGEX vectors, pEZZ18, pRIT2T, and pMC1871 (Pharmacia, Inc.), pKK233-2 and pKK388-1 (Clontech, Inc.), and pProEx-HT (Life Technologies, Inc.) and variants and derivatives thereof. Destination Vectors can also be made from eukaryotic Expression Vectors such as pFastBac, pFastBac HT, pFastBac DUAL, pSFV, and pTet-Splice (Life Technologies, Inc.), pEUK-C1, pPUR, pMAM, pMAMneo, pBI101, pBI121, pDR2, pCMVEBNA, and pYACneo (Clontech), pSVK3, pSVL, pMSG, pCH110, and pKK232-8 (Pharmacia, Inc.), p3'SS, pXT1, pSG5, pPbac, pMbac, pMC1neo, and pOG44 (Stratagene, Inc.), and pYES2, pAC360, pBlueBacHis A,

B, and C, pVL1392, pBsueBacIII, pCDM8, pcDNA1, pZeoSV, pcDNA3 pREP4, pCEP4, and pEBVHis (Invitrogen, Inc.) and variants or derivatives thereof.

Other vectors of particular interest include pUC18, pUC19, pBlueScript, pSPORT, cosmid, phagemids, YACs (yeast artificial chromosomes), BACs (bacterial artificial chromosomes), MACs (mammalian artificial chromosomes), pQE70, pQE60, pQE9 (Quiagen), pBS vectors, PhageScript vectors, BlueScript vectors, pNH8A, pNH16A, pNH18A, pNH46A (Stratagene), pcDNA3 (Invitrogen), pGEX, pTrsfus, pTrc99A, pET-5, pET-9, pKK223-3, pKK233-3, pDR540, pRIT5 (Pharmacia), pSPORT1, pSPORT2, pCMVSPORT2.0 and pSV-SPORT1 (Life Technologies, Inc.) and variants or derivatives thereof.

Additional vectors of interest include pTrxFus, pThioHis, pLEX, pTrcHis, pTrcHis2, pRSET, pBlueBacHis2, pcDNA3.1/His, pcDNA3.1(-)/Myc-His, pSecTag, pEBVHis, pPIC9K, pPIC3.5K, pAO815, pPICZ, pPICZ α , pGAPZ, pGAPZ α , pBlueBac4.5, pBlueBacHis2, pMelBac, pSinRep5, pSinHis, pIND, pIND(SP1), pVgRXR, pcDNA2.1, pYES2, pZER01.1, pZER0-2.1, pCR-Blunt, pSE280, pSE380, pSE420, pVL1392, pVL1393, pCDM8, pcDNA1.1, pcDNA1.1/Amp, pcDNA3.1, pcDNA3.1/Zeo, pSe,SV2, pRc/CMV2, pRc/RSV, pREP4, pREP7, pREP8, pREP9, pREP10, pCEP4, pEBVHis, pCR3.1, pCR2.1, pCR3.1-Uni, and pCRBac from Invitrogen; λ ExCell, λ gt11, pTrc99A, pKK223-3, pGEX-1 λ T, pGEX-2T, pGEX-2TK, pGEX-4T-1, pGEX-4T-2, pGEX-4T-3, pGEX-3X, pGEX-5X-1, pGEX-5X-2, pGEX-5X-3, pEZZ18, pRIT2T, pMC1871, pSVK3, pSVL, pMSG, pCH110, pKK232-8, pSL1180, pNEO, and pUC4K from Pharmacia; pSCREEN-1b(+), pT7Blue(R), pT7Blue-2, pCITE-4abc(+), pOCUS-2, pTag, pET-32 LIC, pET-30 LIC, pBAC-2cp LIC, pBACgus-2cp LIC, pT7Blue-2 LIC, pT7Blue-2, λ SCREEN-1, λ BlueSTAR, pET-3abcd, pET-7abc, pET9abcd, pET11abcd, pET12abc, pET-14b, pET-15b, pET-16b, pET-17b-pET-17xb, pET-19b, pET-20b(+), pET-21abcd(+), pET-22b(+), pET-23abcd(+), pET-24abcd(+), pET-25b(+), pET-26b(+), pET-27b(+), pET-28abc(+), pET-29abc(+), pET-30abc(+), pET-31b(+), pET-32abc(+), pET-33b(+), pBAC-1, pBACgus-1, pBAC4x-1, pBACgus4x-1, pBAC-3cp, pBACgus-2cp, pBACsurf-1, plg, Signal plg, pYX, Selecta Vecta-Neo, Selecta Vecta - Hyg, and Selecta Vecta - Gpt from Novagen; pLexA, pB42AD, pGBT9, pAS2-1,

pGAD424, pACT2, pGAD GL, pGAD GH, pGAD10, pGilda, pEZM3, pEGFP, pEGFP-1, pEGFP-N, pEGFP-C, pEBFP, pGFPuv, pGFP, p6xHis-GFP, pSEAP2-Basic, pSEAP2-Contral, pSEAP2-Promoter, pSEAP2-Enhancer, p β gal-Basic, p β gal-Control, p β gal-Promoter, p β gal-Enhancer, pCMV β , pTet-Off, pTet-On, pTK-Hyg, pRetro-Off, pRetro-On, pIRES1neo, pIRES1hyg, pLXSN, pLNCX, pLAPSN, pMAMneo, pMAMneo-CAT, pMAMneo-LUC, pPUR, pSV2neo, pYEX 4T-1/2/3, pYEX-S1, pBacPAK-His, pBacPAK8/9, pAcUW31, BacPAK6, pTriplEx, λ gt10, λ gt11, pWE15, and λ TriplEx from Clontech; Lambda ZAP II, pBK-CMV, pBK-RSV, pBluescript II KS +/-, pBluescript II SK +/-, pAD-GAL4, pBD-GAL4 Cam, pSurfscrip, Lambda FIX II, Lambda DASH, Lambda EMBL3, Lambda EMBL4, SuperCos, pCR-Script Amp, pCR-Script Cam, pCR-Script Direct, pBS +/-, pBC KS +/-, pBC SK +/-, Phagescript, pCAL-n-EK, pCAL-n, pCAL-c, pCAL-kc, pET-3abcd, pET-11abcd, pSPUTK, pESP-1, pCMVLacI, pOPRSVI/MCS, pOPI3 CAT, pXT1, pSG5, pPbac, pMbac, pMC1neo, pMC1neo Poly A, pOG44, pOG45, pFRT β GAL, pNEO β GAL, pRS403, pRS404, pRS405, pRS406, pRS413, pRS414, pRS415, and pRS416 from Stratagene.

Two-hybrid and reverse two-hybrid vectors of particular interest include pPC86, pDBLeu, pDBTrp, pPC97, p2.5, pGAD1-3, pGAD10, pACT, pACT2, pGADGL, pGADGH, pAS2-1, pGAD424, pGBT8, pGBT9, pGAD-GAL4, pLexA, pBD-GAL4, pHISi, pHISi-1, placZi, pB42AD, pDG202, pJK202, pJG4-5, pNLexA, pYESTrp and variants or derivatives thereof.

Yeast Expression Vectors of particular interest include pESP-1, pESP-2, pESC-His, pESC-Trp, pESC-URA, pESC-Leu (Stratagene), pRS401, pRS402, pRS411, pRS412, pRS421, pRS422, and variants or derivatives thereof.

According to the invention, the vectors comprising one or more nucleic acid molecules encoding one or more recombination sites, or mutants, variants, fragments, or derivatives thereof, may be produced by one of ordinary skill in the art without resorting to undue experimentation using standard molecular biology methods. For example, the vectors of the invention may be produced by introducing one or more of the nucleic acid molecules encoding one or more recombination sites (or mutants, fragments, variants or derivatives thereof) into one or more of the vectors described herein, according to the methods described,

for example, in Maniatis *et al.*, *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York (1982). In a related aspect of the invention, the vectors may be engineered to contain, in addition to one or more nucleic acid molecules encoding one or more recombination sites (or portions thereof), one or more additional physical or functional nucleotide sequences, such as those encoding one or more multiple cloning sites, one or more transcription termination sites, one or more transcriptional regulatory sequences (*e.g.*, one or more promoters, enhancers, or repressors), one or more selection markers or modules, one or more genes or portions of genes encoding a protein or polypeptide of interest, one or more translational signal sequences, one or more nucleotide sequences encoding a fusion partner protein or peptide (*e.g.*, GST, His₆ or thioredoxin), one or more origins of replication, and one or more 5' or 3' polynucleotide tails (particularly a poly-G tail). According to this aspect of the invention, the one or more recombination site nucleotide sequences (or portions thereof) may optionally be operably linked to the one or more additional physical or functional nucleotide sequences described herein.

Preferred vectors according to this aspect of the invention include, but are not limited to: pENTR1A (Figures 10A and 10B), pENTR2B (Figures 11A and 11B), pENTR3C (Figures 12A and 12B), pENTR4 (Figures 13A and 13B), pENTR5 (Figures 14A and 14B), pENTR6 (Figures 15A and 15B), pENTR7 (Figures 16A and 16B), pENTR8 (Figures 17A and 17B), pENTR9 (Figures 18A and 18B), pENTR10 (Figures 19A and 19B), pENTR11 (Figures 20A and 20B), pDEST1 (Figures 21A-D), pDEST2 (Figure 22A-D), pDEST3 (Figure 23A-D), pDEST4 (Figure 24A-D), pDEST5 (Figure 25A-D), pDEST6 (Figure 26A-D), pDEST7 (Figure 27A-C), pDEST8 (Figure 28A-D), pDEST9 (Figure 29A-E), pDEST10 (Figure 30A-D), pDEST11 (Figure 31A-D), pDEST12.2 (also known as pDEST12) (Figure 32A-D), pDEST13 (Figure 33A-C), pDEST14 (Figure 34A-D), pDEST15 (Figure 35A-D), pDEST16 (Figure 36A-D), pDEST17 (Figure 37A-D), pDEST18 (Figure 38A-D), pDEST19 (Figure 39A-D), pDEST20 (Figure 40A-D), pDEST21 (Figure 41A-E), pDEST22 (Figure 42A-D), pDEST23 (Figure 43A-D), pDEST24 (Figure 44A-D), pDEST25 (Figure 45A-D), pDEST26 (Figure 46A-D), pDEST27 (Figure 47A-D), pEXP501 (also known

as pCMVSPORT6) (Figure 48A-B), pDONR201 (also known as pENTR21 attP vector or pAttPkan Donor Vector) (Figure 49), pDONR202 (Figure 50), pDONR203 (also known as pEZ15812) (Figure 51), pDONR204 (Figure 52), pDONR205 (Figure 53), pDONR206 (also known as pENTR22 attP vector or pAttPgen Donor Vector) (Figure 54), pMAB58 (Figure 87), pMAB62 (Figure 88), pDEST28 (Figure 90), pDEST29 (Figure 91), pDEST30 (Figure 92), pDEST31 (Figure 93), pDEST32 (Figure 94), pDEST33 (Figure 95), pDEST34 (Figure 96), pDONR207 (Figure 97), pMAB85 (Figure 98), pMAB86 (Figure 99), and fragments, mutants, variants, and derivatives thereof. However, it will be understood by one of ordinary skill that the present invention also encompasses other vectors not specifically designated herein, which comprise one or more of the isolated nucleic acid molecules of the invention encoding one or more recombination sites or portions thereof (or mutants, fragments, variants or derivatives thereof), and which may further comprise one or more additional physical or functional nucleotide sequences described herein which may optionally be operably linked to the one or more nucleic acid molecules encoding one or more recombination sites or portions thereof. Such additional vectors may be produced by one of ordinary skill according to the guidance provided in the present specification.

Polymerases

Preferred polypeptides having reverse transcriptase activity (*i.e.*, those polypeptides able to catalyze the synthesis of a DNA molecule from an RNA template) for use in accordance with the present invention include, but are not limited to Moloney Murine Leukemia Virus (M-MLV) reverse transcriptase, Rous Sarcoma Virus (RSV) reverse transcriptase, Avian Myeloblastosis Virus (AMV) reverse transcriptase, Rous Associated Virus (RAV) reverse transcriptase, Myeloblastosis Associated Virus (MAV) reverse transcriptase, Human Immunodeficiency Virus (HIV) reverse transcriptase, retroviral reverse transcriptase, retrotransposon reverse transcriptase, hepatitis B reverse transcriptase, cauliflower mosaic virus reverse transcriptase and bacterial reverse transcriptase. Particularly preferred are those polypeptides having reverse

transcriptase activity that are also substantially reduced in RNase H activity (*i.e.*, "RNase H" polypeptides). By a polypeptide that is "substantially reduced in RNase H activity" is meant that the polypeptide has less than about 20%, more preferably less than about 15%, 10% or 5%, and most preferably less than about 2%, of the RNase H activity of a wildtype or RNase H⁺ enzyme such as wildtype M-MLV reverse transcriptase. The RNase H activity may be determined by a variety of assays, such as those described, for example, in U.S. Patent No. 5,244,797, in Kotewicz, M.L. *et al.*, *Nucl. Acids Res.* 16:265 (1988) and in Gerard, G.F., *et al.*, *FOCUS* 14(5):91 (1992), the disclosures of all of which are fully incorporated herein by reference. Suitable RNase H⁺ polypeptides for use in the present invention include, but are not limited to, M-MLV H⁺ reverse transcriptase, RSV H⁺ reverse transcriptase, AMV H⁺ reverse transcriptase, RAV H⁺ reverse transcriptase, MAV H⁺ reverse transcriptase, HIV H⁺ reverse transcriptase, THERMOSCRIPT™ reverse transcriptase and THERMOSCRIPT™ II reverse transcriptase, and SUPERSCRIPT™ I reverse transcriptase and SUPERSCRIPT™ II reverse transcriptase, which are obtainable, for example, from Life Technologies, Inc. (Rockville, Maryland). See generally published PCT application WO 98/47912.

Other polypeptides having nucleic acid polymerase activity suitable for use in the present methods include thermophilic DNA polymerases such as DNA polymerase I, DNA polymerase III, Klenow fragment, T7 polymerase, and T5 polymerase, and thermostable DNA polymerases including, but not limited to, *Thermus thermophilus* (*Tth*) DNA polymerase, *Thermus aquaticus* (*Taq*) DNA polymerase, *Thermotoga neopolitana* (*Tne*) DNA polymerase, *Thermotoga maritima* (*Tma*) DNA polymerase, *Thermococcus litoralis* (*Tli* or VENT®) DNA polymerase, *Pyrococcus furiosus* (*Pfu*) DNA polymerase, *Pyrococcus* species GB-D (or DEEPVENT®) DNA polymerase, *Pyrococcus woosii* (*Pwo*) DNA polymerase, *Bacillus sterothermophilus* (*Bst*) DNA polymerase, *Sulfolobus acidocaldarius* (*Sac*) DNA polymerase, *Thermoplasma acidophilum* (*Tac*) DNA polymerase, *Thermus flavus* (*Tfl/Tub*) DNA polymerase, *Thermus ruber* (*Tru*) DNA polymerase, *Thermus brockianus* (DYNAZYME®) DNA polymerase, *Methanobacterium thermoautotrophicum* (*Mth*) DNA polymerase, and mutants,

variants and derivatives thereof. Such polypeptides are available commercially, for example from Life Technologies, Inc. (Rockville, MD), New England BioLabs (Beverly, MA), and Sigma/Aldrich (St. Louis, MO).

Host Cells

The invention also relates to host cells comprising one or more of the nucleic acid molecules or vectors of the invention, particularly those nucleic acid molecules and vectors described in detail herein. Representative host cells that may be used according to this aspect of the invention include, but are not limited to, bacterial cells, yeast cells, plant cells and animal cells. Preferred bacterial host cells include *Escherichia* spp. cells (particularly *E. coli* cells and most particularly *E. coli* strains DH10B, Stbl2, DH5 α , DB3, DB3.1 (preferably *E. coli* LIBRARY EFFICIENCY \circledR DB3.1TM Competent Cells; Life Technologies, Inc., Rockville, MD), DB4 and DB5; see U.S. Provisional Application No. 60/122,392, filed on March 2, 1999, the disclosure of which is incorporated by reference herein in its entirety), *Bacillus* spp. cells (particularly *B. subtilis* and *B. megaterium* cells), *Streptomyces* spp. cells, *Erwinia* spp. cells, *Klebsiella* spp. cells, *Serratia* spp. cells (particularly *S. marcessans* cells), *Pseudomonas* spp. cells (particularly *P. aeruginosa* cells), and *Salmonella* spp. cells (particularly *S. typhimurium* and *S. typhi* cells). Preferred animal host cells include insect cells (most particularly *Drosophila melanogaster* cells, *Spodoptera frugiperda* Sf9 and Sf21 cells and *Trichoplusia* High-Five cells), nematode cells (particularly *C. elegans* cells), avian cells, amphibian cells (particularly *Xenopus laevis* cells), reptilian cells, and mammalian cells (most particularly CHO, COS, VERO, BHK and human cells). Preferred yeast host cells include *Saccharomyces cerevisiae* cells and *Pichia pastoris* cells. These and other suitable host cells are available commercially, for example from Life Technologies, Inc. (Rockville, Maryland), American Type Culture Collection (Manassas, Virginia), and Agricultural Research Culture Collection (NRRL; Peoria, Illinois).

Methods for introducing the nucleic acid molecules and/or vectors of the invention into the host cells described herein, to produce host cells comprising one or more of the nucleic acid molecules and/or vectors of the invention, will be

familiar to those of ordinary skill in the art. For instance, the nucleic acid molecules and/or vectors of the invention may be introduced into host cells using well known techniques of infection, transduction, transfection, and transformation. The nucleic acid molecules and/or vectors of the invention may be introduced alone or in conjunction with other the nucleic acid molecules and/or vectors. Alternatively, the nucleic acid molecules and/or vectors of the invention may be introduced into host cells as a precipitate, such as a calcium phosphate precipitate, or in a complex with a lipid. Electroporation also may be used to introduce the nucleic acid molecules and/or vectors of the invention into a host. Likewise, such molecules may be introduced into chemically competent cells such as *E. coli*. If the vector is a virus, it may be packaged *in vitro* or introduced into a packaging cell and the packaged virus may be transduced into cells. Hence, a wide variety of techniques suitable for introducing the nucleic acid molecules and/or vectors of the invention into cells in accordance with this aspect of the invention are well known and routine to those of skill in the art. Such techniques are reviewed at length, for example, in Sambrook, J., *et al.*, *Molecular Cloning, a Laboratory Manual, 2nd Ed.*, Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press, pp. 16.30-16.55 (1989), Watson, J.D., *et al.*, *Recombinant DNA, 2nd Ed.*, New York: W.H. Freeman and Co., pp. 213-234 (1992), and Winnacker, E.-L., *From Genes to Clones*, New York: VCH Publishers (1987), which are illustrative of the many laboratory manuals that detail these techniques and which are incorporated by reference herein in their entireties for their relevant disclosures.

Polypeptides

In another aspect, the invention relates to polypeptides encoded by the nucleic acid molecules of the invention (including polypeptides and amino acid sequences encoded by all possible reading frames of the nucleic acid molecules of the invention), and to methods of producing such polypeptides. Polypeptides of the present invention include purified or isolated natural products, products of chemical synthetic procedures, and products produced by recombinant techniques from a prokaryotic or eukaryotic host, including, for example, bacterial, yeast, insect, mammalian, avian and higher plant cells.

The polypeptides of the invention may be produced by synthetic organic chemistry, and are preferably produced by standard recombinant methods, employing one or more of the host cells of the invention comprising the vectors or isolated nucleic acid molecules of the invention. According to the invention, polypeptides are produced by cultivating the host cells of the invention (which comprise one or more of the nucleic acid molecules of the invention, preferably contained within an Expression Vector) under conditions favoring the expression of the nucleotide sequence contained on the nucleic acid molecule of the invention, such that the polypeptide encoded by the nucleic acid molecule of the invention is produced by the host cell. As used herein, "conditions favoring the expression of the nucleotide sequence" or "conditions favoring the production of a polypeptide" include optimal physical (e.g., temperature, humidity, etc.) and nutritional (e.g., culture medium, ionic) conditions required for production of a recombinant polypeptide by a given host cell. Such optimal conditions for a variety of host cells, including prokaryotic (bacterial), mammalian, insect, yeast, and plant cells will be familiar to one of ordinary skill in the art, and may be found, for example, in Sambrook, J., *et al.*, *Molecular Cloning, A Laboratory Manual, 2nd Ed.*, Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press, (1989), Watson, J.D., *et al.*, *Recombinant DNA, 2nd Ed.*, New York: W.H. Freeman and Co., and Winnacker, E.-L., *From Genes to Clones*, New York: VCH Publishers (1987).

In some aspects, it may be desirable to isolate or purify the polypeptides of the invention (e.g., for production of antibodies as described below), resulting in the production of the polypeptides of the invention in isolated form. The polypeptides of the invention can be recovered and purified from recombinant cell cultures by well-known methods of protein purification that are routine in the art, including ammonium sulfate or ethanol precipitation, acid extraction, anion or cation exchange chromatography, phosphocellulose chromatography, hydrophobic interaction chromatography, affinity chromatography, hydroxylapatite chromatography and lectin chromatography. For example, His6 or GST fusion tags on polypeptides made by the methods of the invention may be isolated using appropriate affinity chromatography matrices which bind polypeptides bearing

His6 or GST tags, as will be familiar to one of ordinary skill in the art. Polypeptides of the present invention include naturally purified products, products of chemical synthetic procedures, and products produced by recombinant techniques from a prokaryotic or eukaryotic host, including, for example, bacterial, yeast, higher plant, insect and mammalian cells. Depending upon the host employed in a recombinant production procedure, the polypeptides of the present invention may be glycosylated or may be non-glycosylated. In addition, polypeptides of the invention may also include an initial modified methionine residue, in some cases as a result of host-mediated processes.

Isolated polypeptides of the invention include those comprising the amino acid sequences encoded by one or more of the reading frames of the polynucleotides comprising one or more of the recombination site-encoding nucleic acid molecules of the invention, including those encoding *attB1*, *attB2*, *attP1*, *attP2*, *attL1*, *attL2*, *attR1* and *attR2* having the nucleotide sequences set forth in Figure 9 (or nucleotide sequences complementary thereto), or fragments, variants, mutants and derivatives thereof; the complete amino acid sequences encoded by the polynucleotides contained in the deposited clones described herein; the amino acid sequences encoded by polynucleotides which hybridize under stringent hybridization conditions to polynucleotides having the nucleotide sequences encoding the recombination site sequences of the invention as set forth in Figure 9 (or a nucleotide sequence complementary thereto); or a peptide or polypeptide comprising a portion or a fragment of the above polypeptides. The invention also relates to additional polypeptides having one or more additional amino acids linked (typically by peptidyl bonds to form a nascent polypeptide) to the polypeptides encoded by the recombination site nucleotide sequences or the deposited clones. Such additional amino acid residues may comprise one or more functional peptide sequences, for example one or more fusion partner peptides (e.g., GST, His₆, Trx, etc.) and the like.

As used herein, the terms "protein," "peptide," "oligopeptide" and "polypeptide" are considered synonymous (as is commonly recognized) and each term can be used interchangeably as the context requires to indicate a chain of two or more amino acids, preferably five or more amino acids, or more preferably ten

or more amino acids, coupled by (a) peptidyl linkage(s), unless otherwise defined in the specific contexts below. As is commonly recognized in the art, all polypeptide formulas or sequences herein are written from left to right and in the direction from amino terminus to carboxy terminus.

5 It will be recognized by those of ordinary skill in the art that some amino acid sequences of the polypeptides of the invention can be varied without significant effect on the structure or function of the polypeptides. If such differences in sequence are contemplated, it should be remembered that there will be critical areas on the protein which determine structure and activity. In general, it is possible to replace residues which form the tertiary structure, provided that
10 residues performing a similar function are used. In other instances, the type of residue may be completely unimportant if the alteration occurs at a non-critical region of the polypeptide.

Thus, the invention further includes variants of the polypeptides of the invention, including allelic variants, which show substantial structural homology
15 to the polypeptides described herein, or which include specific regions of these polypeptides such as the portions discussed below. Such mutants may include deletions, insertions, inversions, repeats, and type substitutions (for example, substituting one hydrophilic residue for another, but not strongly hydrophilic for
20 strongly hydrophobic as a rule). Small changes or such "neutral" or "conservative" amino acid substitutions will generally have little effect on activity.

Typical conservative substitutions are the replacements, one for another, among the aliphatic amino acids Ala, Val, Leu and Ile; interchange of the hydroxylated residues Ser and Thr; exchange of the acidic residues Asp and Glu;
25 substitution between the amidated residues Asn and Gln; exchange of the basic residues Lys and Arg; and replacements among the aromatic residues Phe and Tyr.

Thus, the fragment, derivative or analog of the polypeptides of the invention, such as those comprising peptides encoded by the recombination site nucleotide sequences described herein, may be (i) one in which one or more of the amino acid
30 residues are substituted with a conservative or non-conservative amino acid residue (preferably a conservative amino acid residue), and such substituted amino acid residue may be encoded by the genetic code or may be an amino acid (*e.g.*,

desmosine, citrulline, ornithine, etc.) that is not encoded by the genetic code; (ii) one in which one or more of the amino acid residues includes a substituent group (e.g., a phosphate, hydroxyl, sulfate or other group) in addition to the normal "R" group of the amino acid; (iii) one in which the mature polypeptide is fused with another compound, such as a compound to increase the half-life of the polypeptide (for example, polyethylene glycol), or (iv) one in which additional amino acids are fused to the mature polypeptide, such as an immunoglobulin Fc region peptide, a leader or secretory sequence, a sequence which is employed for purification of the mature polypeptide (such as GST) or a proprotein sequence. Such fragments, derivatives and analogs are intended to be encompassed by the present invention, and are within the scope of those skilled in the art from the teachings herein and the state of the art at the time of invention.

The polypeptides of the present invention are preferably provided in an isolated form, and preferably are substantially purified. Recombinantly produced versions of the polypeptides of the invention can be substantially purified by the one-step method described in Smith and Johnson, *Gene* 67:31-40 (1988). As used herein, the term "substantially purified" means a preparation of an individual polypeptide of the invention wherein at least 50%, preferably at least 60%, 70%, or 75% and more preferably at least 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% (by mass) of contaminating proteins (*i.e.*, those that are not the individual polypeptides described herein or fragments, variants, mutants or derivatives thereof) have been removed from the preparation.

The polypeptides of the present invention include those which are at least about 50% identical, at least 60% identical, at least 65% identical, more preferably at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98% or at least about 99% identical, to the polypeptides described herein. For example, preferred *attB1*-containing polypeptides of the invention include those that are at least about 50% identical, at least 60% identical, at least 65% identical, more preferably at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98% or at least about 99% identical,

to the polypeptide(s) encoded by the three reading frames of a polynucleotide comprising a nucleotide sequence of *attB1* having a nucleic acid sequence as set forth in Figure 9 (or a nucleic acid sequence complementary thereto), to a polypeptide encoded by a polynucleotide contained in the deposited cDNA clones described herein, or to a polypeptide encoded by a polynucleotide hybridizing under stringent conditions to a polynucleotide comprising a nucleotide sequence of *attB1* having a nucleic acid sequence as set forth in Figure 9 (or a nucleic acid sequence complementary thereto). Analogous polypeptides may be prepared that are at least about 65% identical, more preferably at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98% or at least about 99% identical, to the *attB2*, *attP1*, *attP2*, *attL1*, *attL2*, *attR1* and *attR2* polypeptides of the invention as depicted in Figure 9. The present polypeptides also include portions or fragments of the above-described polypeptides with at least 5, 10, 15, 20, or 25 amino acids.

By a polypeptide having an amino acid sequence at least, for example, 65% "identical" to a reference amino acid sequence of a given polypeptide of the invention is intended that the amino acid sequence of the polypeptide is identical to the reference sequence except that the polypeptide sequence may include up to 35 amino acid alterations per each 100 amino acids of the reference amino acid sequence of a given polypeptide of the invention. In other words, to obtain a polypeptide having an amino acid sequence at least 65% identical to a reference amino acid sequence, up to 35% of the amino acid residues in the reference sequence may be deleted or substituted with another amino acid, or a number of amino acids up to 35% of the total amino acid residues in the reference sequence may be inserted into the reference sequence. These alterations of the reference sequence may occur at the amino (N-) or carboxy (C-) terminal positions of the reference amino acid sequence or anywhere between those terminal positions, interspersed either individually among residues in the reference sequence or in one or more contiguous groups within the reference sequence. As a practical matter, whether a given amino acid sequence is, for example, at least 65% identical to the amino acid sequence of a given polypeptide of the invention can be determined

conventionally using known computer programs such as those described above for nucleic acid sequence identity determinations, or more preferably using the CLUSTAL W program (Thompson, J.D., *et al.*, *Nucleic Acids Res.* 22:4673-4680 (1994)).

5 The polypeptides of the present invention can be used as molecular weight markers on SDS-PAGE gels or on molecular sieve gel filtration columns using methods well known to those of skill in the art. In addition, as described in detail below, the polypeptides of the present invention can be used to raise polyclonal and monoclonal antibodies which are useful in a variety of assays for detecting
10 protein expression, localization, detection of interactions with other molecules, or for the isolation of a polypeptide (including a fusion polypeptide) of the invention.

 In another aspect, the present invention provides a peptide or polypeptide comprising an epitope-bearing portion of a polypeptide of the invention, which may be used to raise antibodies, particularly monoclonal antibodies, that bind
15 specifically to a one or more of the polypeptides of the invention. The epitope of this polypeptide portion is an immunogenic or antigenic epitope of a polypeptide of the invention. An "immunogenic epitope" is defined as a part of a protein that elicits an antibody response when the whole protein is the immunogen. These immunogenic epitopes are believed to be confined to a few loci on the molecule.
20 On the other hand, a region of a protein molecule to which an antibody can bind is defined as an "antigenic epitope." The number of immunogenic epitopes of a protein generally is less than the number of antigenic epitopes (*see, e.g.*, Geysen *et al.*, *Proc. Natl. Acad. Sci. USA* 81:3998-4002 (1983)).

 As to the selection of peptides or polypeptides bearing an antigenic epitope
25 (*i.e.*, that contain a region of a protein molecule to which an antibody can bind), it is well-known in the art that relatively short synthetic peptides that mimic part of a protein sequence are routinely capable of eliciting an antiserum that reacts with the partially mimicked protein (*see, e.g.*, Sutcliffe, J.G., *et al.*, *Science* 219:660-666 (1983)). Peptides capable of eliciting protein-reactive sera are
30 frequently represented in the primary sequence of a protein, can be characterized by a set of simple chemical rules, and are not confined to the immunodominant regions of intact proteins (*i.e.*, immunogenic epitopes) or to the amino or carboxy

termini. Peptides that are extremely hydrophobic and those of six or fewer residues generally are ineffective at inducing antibodies that bind to the mimicked protein; longer peptides, especially those containing proline residues, usually are effective (Sutcliffe, J.G., *et al.*, *Science* 219:660-666 (1983)).

5 Epitope-bearing peptides and polypeptides of the invention designed according to the above guidelines preferably contain a sequence of at least five, more preferably at least seven or more amino acids contained within the amino acid sequence of a polypeptide of the invention. However, peptides or polypeptides comprising a larger portion of an amino acid sequence of a polypeptide of the
10 invention, containing about 30 to about 50 amino acids, or any length up to and including the entire amino acid sequence of a given polypeptide of the invention, also are considered epitope-bearing peptides or polypeptides of the invention and also are useful for inducing antibodies that react with the mimicked protein. Preferably, the amino acid sequence of the epitope-bearing peptide is selected to
15 provide substantial solubility in aqueous solvents (*i.e.*, the sequence includes relatively hydrophilic residues and highly hydrophobic sequences are preferably avoided); sequences containing proline residues are particularly preferred.

Non-limiting examples of epitope-bearing polypeptides or peptides that can be used to generate antibodies specific for the polypeptides of the invention include
20 certain epitope-bearing regions of the polypeptides comprising amino acid sequences encoded by polynucleotides comprising one or more of the recombination site-encoding nucleic acid molecules of the invention, including those encoding *attB1*, *attB2*, *attP1*, *attP2*, *attL1*, *attL2*, *attR1* and *attR2* having the nucleotide sequences set forth in Figure 9 (or a nucleotide sequence
25 complementary thereto); the complete amino acid sequences encoded by the three reading frames of the polynucleotides contained in the deposited clones described herein; and the amino acid sequences encoded by all reading frames of polynucleotides which hybridize under stringent hybridization conditions to polynucleotides having the nucleotide sequences encoding the recombination site
30 sequences (or portions thereof) of the invention as set forth in Figure 9 (or a nucleic acid sequence complementary thereto). Other epitope-bearing polypeptides or peptides that may be used to generate antibodies specific for the polypeptides

of the invention will be apparent to one of ordinary skill in the art based on the primary amino acid sequences of the polypeptides of the invention described herein, via the construction of Kyte-Doolittle hydrophilicity and Jameson-Wolf antigenic index plots of the polypeptides of the invention using, for example, 5 PROTEAN computer software (DNASTAR, Inc.; Madison, Wisconsin).

The epitope-bearing peptides and polypeptides of the invention may be produced by any conventional means for making peptides or polypeptides including recombinant means using nucleic acid molecules of the invention. For instance, a short epitope-bearing amino acid sequence may be fused to a larger 10 polypeptide which acts as a carrier during recombinant production and purification, as well as during immunization to produce anti-peptide antibodies. Epitope-bearing peptides also may be synthesized using known methods of chemical synthesis (*see, e.g.*, U.S. Patent No. 4,631,211 and Houghten, R. A., *Proc. Natl. Acad. Sci. USA* 82:5131-5135 (1985), both of which are incorporated 15 by reference herein in their entireties).

As one of skill in the art will appreciate, the polypeptides of the present invention and epitope-bearing fragments thereof may be immobilized onto a solid support, by techniques that are well-known and routine in the art. By "solid support" is intended any solid support to which a peptide can be immobilized. 20 Such solid supports include, but are not limited to nitrocellulose, diazocellulose, glass, polystyrene, polyvinylchloride, polypropylene, polyethylene, dextran, Sepharose, agar, starch, nylon, beads and microtitre plates. Linkage of the peptide of the invention to a solid support can be accomplished by attaching one or both ends of the peptide to the support. Attachment may also be made at one or more 25 internal sites in the peptide. Multiple attachments (both internal and at the ends of the peptide) may also be used according to the invention. Attachment can be via an amino acid linkage group such as a primary amino group, a carboxyl group, or a sulfhydryl (SH) group or by chemical linkage groups such as with cyanogen bromide (CNBr) linkage through a spacer. For non-covalent attachments to the support, addition of an affinity tag sequence to the peptide can be used such as 30 GST (Smith, D.B., and Johnson, K.S., *Gene* 67:31 (1988)), polyhistidines (Hochuli, E., *et al.*, *J. Chromatog.* 411:77 (1987)), or biotin. Such affinity tags

may be used for the reversible attachment of the peptide to the support. Such immobilized polypeptides or fragments may be useful, for example, in isolating antibodies directed against one or more of the polypeptides of the invention, or other proteins or peptides that recognize other proteins or peptides that bind to one or more of the polypeptides of the invention, as described below.

As one of skill in the art will also appreciate, the polypeptides of the present invention and the epitope-bearing fragments thereof described herein can be combined with one or more fusion partner proteins or peptides, or portions thereof, including but not limited to GST, His₆, Trx, and portions of the constant domain of immunoglobulins (Ig), resulting in chimeric or fusion polypeptides. These fusion polypeptides facilitate purification of the polypeptides of the invention (EP 0 394 827; Traunecker *et al.*, *Nature* 331:84-86 (1988)) for use in analytical or diagnostic (including high-throughput) format.

Antibodies

In another aspect, the invention relates to antibodies that recognize and bind to the polypeptides (or epitope-bearing fragments thereof) or nucleic acid molecules (or portions thereof) of the invention. In a related aspect, the invention relates to antibodies that recognize and bind to one or more polypeptides encoded by all reading frames of one or more recombination site nucleic acid sequences or portions thereof, or to one or more nucleic acid molecules comprising one or more recombination site nucleic acid sequences or portions thereof, including but not limited to *att* sites (including *attB*1, *attB*2, *attP*1, *attP*2, *attL*1, *attL*2, *attR*1, *attR*2 and the like), *lox* sites (*e.g.*, *loxP*, *loxP*511, and the like), FRT, and the like, or mutants, fragments, variants and derivatives thereof. See generally U.S. Patent No. 5,888,732, which is incorporated herein by reference in its entirety. The antibodies of the present invention may be polyclonal or monoclonal, and may be prepared by any of a variety of methods and in a variety of species according to methods that are well-known in the art. See, for instance, U.S. Patent No. 5,587,287; Sutcliffe, J.G., *et al.*, *Science* 219:660-666 (1983); Wilson *et al.*, *Cell* 37: 767 (1984); and Bittle, F.J., *et al.*, *J. Gen. Virol.* 66:2347-2354 (1985). Antibodies specific for any of the polypeptides or nucleic acid molecules described

herein, such as antibodies specifically binding to one or more of the polypeptides encoded by the recombination site nucleotide sequences, or one or more nucleic acid molecules, described herein or contained in the deposited clones, antibodies against fusion polypeptides (*e.g.*, binding to fusion polypeptides between one or more of the fusion partner proteins and one or more of the recombination site polypeptides of the invention, as described herein), and the like, can be raised against the intact polypeptides or polynucleotides of the invention or one or more antigenic polypeptide fragments thereof.

As used herein, the term "antibody" (Ab) may be used interchangeably with the terms "polyclonal antibody" or "monoclonal antibody" (mAb), except in specific contexts as described below. These terms, as used herein, are meant to include intact molecules as well as antibody fragments (such as, for example, Fab and F(ab')₂ fragments) which are capable of specifically binding to a polypeptide or nucleic acid molecule of the invention or a portion thereof. It will therefore be appreciated that, in addition to the intact antibodies of the invention, Fab, F(ab')₂ and other fragments of the antibodies described herein, and other peptides and peptide fragments that bind one or more polypeptides or polynucleotides of the invention, are also encompassed within the scope of the invention. Such antibody fragments are typically produced by proteolytic cleavage of intact antibodies, using enzymes such as papain (to produce Fab fragments) or pepsin (to produce F(ab')₂ fragments). Antibody fragments, and peptides or peptide fragments, may also be produced through the application of recombinant DNA technology or through synthetic chemistry.

Epitope-bearing peptides and polypeptides, and nucleic acid molecules or portions thereof, of the invention may be used to induce antibodies according to methods well known in the art, as generally described herein (*see, e.g.*, Sutcliffe, *et al.*, *supra*; Wilson, *et al.*, *supra*; and Bittle, F. J., *et al.*, *J. Gen. Virol.* 66:2347-2354 (1985)).

Polyclonal antibodies according to this aspect of the invention may be made by immunizing an animal with one or more of the polypeptides or nucleic acid molecules of the invention described herein or portions thereof according to standard techniques (*see, e.g.*, Harlow, E., and Lane, D., *Antibodies: A*

Laboratory Manual, Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press (1988); Kaufman, P.B., *et al.*, In: *Handbook of Molecular and Cellular Methods in Biology and Medicine*, Boca Raton, Florida: CRC Press, pp. 468-469 (1995)). For producing antibodies that recognize and bind to the polypeptides or nucleic acid molecules of the invention or portions thereof, animals may be immunized with free peptide or free nucleic acid molecules; however, antibody titer may be boosted by coupling of the peptide to a macromolecular carrier, such as albumin, KLH, or tetanus toxoid (particularly for producing antibodies against the nucleic acid molecules of the invention or portions thereof; see Harlow and Lane, *supra*, at page 154), or to a solid phase carrier such as a latex or glass microbead. For instance, peptides containing cysteine may be coupled to carrier using a linker such as m-maleimidobenzoyl-N-hydroxysuccinimide ester (MBS), while other peptides may be coupled to carrier using a more general linking agent such as glutaraldehyde. Animals such as rabbits, rats and mice may be immunized with either free (if the polypeptide immunogen is larger than about 25 amino acids in length) or carrier-coupled peptides or nucleic acid molecules, for instance, by intraperitoneal and/or intradermal injection of emulsions containing about 100 µg peptide, polynucleotide, or carrier protein, and Freund's adjuvant. Several booster injections may be needed, for instance, at intervals of about two weeks, to provide a useful titer of antibody which can be detected, for example, by ELISA assay using free peptide or nucleic acid molecule adsorbed to a solid surface. In another approach, cells expressing one or more of the polypeptides or polynucleotides of the invention or an antigenic fragment thereof can be administered to an animal in order to induce the production of sera containing polyclonal antibodies, according to routine immunological methods. In yet another method, a preparation of one or more of the polypeptides or polynucleotides of the invention is prepared and purified as described herein, to render it substantially free of natural contaminants. Such a preparation may then be introduced into an animal in order to produce polyclonal antisera of greater specific activity. The titer of antibodies in serum from an immunized animal, regardless of the method of immunization used, may be increased by selection of anti-peptide or anti-polynucleotide antibodies, for

instance, by adsorption to the peptide or polynucleotide on a solid support and elution of the selected antibodies according to methods well known in the art.

In an alternative method, the antibodies of the present invention are monoclonal antibodies (or fragments thereof which bind to one or more of the polypeptides of the invention). Such monoclonal antibodies can be prepared using hybridoma technology (Kohler *et al.*, *Nature* 256:495 (1975); Köhler *et al.*, *Eur. J. Immunol.* 6:511 (1976); Köhler *et al.*, *Eur. J. Immunol.* 6:292 (1976); Hammerling *et al.*, In: *Monoclonal Antibodies and T-Cell Hybridomas*, Elsevier, N.Y., pp. 563-681 (1981)). In general, such procedures involve immunizing an animal (preferably a mouse) with a polypeptide or polynucleotide of the invention (or a fragment thereof), or with a cell expressing a polypeptide or polynucleotide of the invention (or a fragment thereof). The splenocytes of such mice are extracted and fused with a suitable myeloma cell line. Any suitable myeloma cell line may be employed in accordance with the present invention; however, it is preferable to employ the parent myeloma cell line (SP₂O), available from the American Type Culture Collection, Rockville, Maryland. After fusion, the resulting hybridoma cells are selectively maintained in HAT medium, and then cloned by limiting dilution as described by Wands *et al.* (*Gastroenterol.* 80:225-232 (1981)). The hybridoma cells obtained through such a selection are then assayed to identify clones which secrete antibodies capable of binding one or more of the polypeptides or nucleic acid molecules of the invention, or fragments thereof. Hence, the present invention also provides hybridoma cells and cell lines producing monoclonal antibodies of the invention, particularly that recognize and bind to one or more of the polypeptides or nucleic acid molecules of the invention.

Alternatively, additional antibodies capable of binding to one or more of the polypeptides of the invention, or fragments thereof, may be produced in a two-step procedure through the use of anti-idiotypic antibodies. Such a method makes use of the fact that antibodies are themselves antigens, and that, therefore, it is possible to obtain an antibody which binds to a second antibody. In accordance with this method, antibodies specific for one or more of the polypeptides or polynucleotides of the invention, prepared as described above, are used to immunize an animal, preferably a mouse. The splenocytes of such an

animal are then used to produce hybridoma cells, and the hybridoma cells are screened to identify clones which produce an antibody whose ability to bind to an antibody specific for one or more of the polypeptides or polynucleotides of the invention can be blocked by polypeptides of the invention themselves. Such antibodies comprise anti-idiotypic antibodies to the antibodies recognizing one or more of the polypeptides or polynucleotides of the invention, and can be used to immunize an animal to induce formation of further antibodies specific for one or more of the polypeptides or polynucleotides of the invention.

For use, the antibodies of the invention may optionally be detectably labeled by covalent or non-covalent attachment of one or more labels, including but not limited to chromogenic, enzymatic, radioisotopic, isotopic, fluorescent, toxic, chemiluminescent, or nuclear magnetic resonance contrast agents or other labels.

Examples of suitable enzyme labels include malate dehydrogenase, staphylococcal nuclease, delta-5-steroid isomerase, yeast-alcohol dehydrogenase, alpha-glycerol phosphate dehydrogenase, triose phosphate isomerase, peroxidase, alkaline phosphatase, asparaginase, glucose oxidase, beta-galactosidase, ribonuclease, urease, catalase, glucose-6-phosphate dehydrogenase, glucoamylase, and acetylcholine esterase.

Examples of suitable radioisotopic labels include ^3H , ^{111}In , ^{125}I , ^{131}I , ^{32}P , ^{35}S , ^{14}C , ^{51}Cr , ^{57}Co , ^{58}Co , ^{59}Fe , ^{75}Se , ^{152}Eu , ^{90}Y , ^{67}Cu , ^{217}Bi , ^{211}At , ^{212}Pb , ^{47}Sc , ^{109}Pd , etc. ^{111}In is a preferred isotope where in vivo imaging is used since it avoids the problem of dehalogenation of the ^{125}I or ^{131}I -labeled monoclonal antibody by the liver. In addition, this radionuclide has a more favorable gamma emission energy for imaging (Perkins *et al.*, *Eur. J. Nucl. Med.* 10:296-301 (1985); Carasquillo *et al.*, *J. Nucl. Med.* 28:281-287 (1987)). For example, ^{111}In coupled to monoclonal antibodies with 1-(P-isothiocyanatobenzyl)-DPTA has shown little uptake in non-tumorous tissues, particularly the liver, and therefore enhances specificity of tumor localization (Esteban *et al.*, *J. Nucl. Med.* 28:861-870 (1987)).

Examples of suitable non-radioactive isotopic labels include ^{157}Gd , ^{55}Mn , ^{162}Dy , ^{52}Tr , and ^{56}Fe .

Examples of suitable fluorescent labels include an ^{152}Eu label, a fluorescein label, an isothiocyanate label, a rhodamine label, a phycoerythrin label, a

phycocyanin label, an allophycocyanin label, an o-phthaldehyde label, a green fluorescent protein (GFP) label, and a fluorescamine label.

Examples of suitable toxin labels include diphtheria toxin, ricin, and cholera toxin.

5 Examples of chemiluminescent labels include a luminal label, an isoluminal label, an aromatic acridinium ester label, an imidazole label, an acridinium salt label, an oxalate ester label, a luciferin label, a luciferase label, and an aequorin label.

10 Examples of nuclear magnetic resonance contrasting agents include heavy metal nuclei such as Gd, Mn, and iron.

Typical techniques for binding the above-described labels to the antibodies of the invention are provided by Kennedy *et al.*, *Clin. Chim. Acta* 70:1-31 (1976), and Schurs *et al.*, *Clin. Chim. Acta* 81:1-40 (1977). Coupling techniques mentioned in the latter are the glutaraldehyde method, the periodate method, the dimaleimide method, the m-maleimidobenzyl-N-hydroxy-succinimide ester method, all of which methods are incorporated by reference herein.

15 It will be appreciated by one of ordinary skill that the antibodies of the present invention may alternatively be coupled to a solid support, to facilitate, for example, chromatographic and other immunological procedures using such solid phase-immobilized antibodies. Included among such procedures are the use of the antibodies of the invention to isolate or purify polypeptides comprising one or more epitopes encoded by the nucleic acid molecules of the invention (which may be fusion polypeptides or other polypeptides of the invention described herein), or to isolate or purify polynucleotides comprising one or more recombination site sequences of the invention or portions thereof. Methods for isolation and purification of polypeptides (and, by analogy, polynucleotides) by affinity chromatography, for example using the antibodies of the invention coupled to a solid phase support, are well-known in the art and will be familiar to one of ordinary skill. The antibodies of the invention may also be used in other applications, for example to cross-link or couple two or more proteins, polypeptides, polynucleotides, or portions thereof into a structural and/or functional complex. In one such use, an antibody of the invention may have two

20

25

30

or more distinct epitope-binding regions that may bind, for example, a first polypeptide (which may be a polypeptide of the invention) at one epitope-binding region on the antibody and a second polypeptide (which may be a polypeptide of the invention) at a second epitope-binding region on the antibody, thereby bringing the first and second polypeptides into close proximity to each other such that the first and second polypeptides are able to interact structurally and/or functionally (as, for example, linking an enzyme and its substrate to carry out enzymatic catalysis, or linking an effector molecule and its receptor to carry out or induce a specific binding of the effector molecule to the receptor or a response to the effector molecule mediated by the receptor). Additional applications for the antibodies of the invention include, for example, the preparation of large-scale arrays of the antibodies, polypeptides, or nucleic acid molecules of the invention, or portions thereof, on a solid support, for example to facilitate high-throughput screening of protein or RNA expression by host cells containing nucleic acid molecules of the invention (known in the art as "chip array" protocols; *see, e.g.*, U.S. Patent Nos. 5,856,101, 5,837,832, 5,770,456, 5,744,305, 5,631,734, and 5,593,839, which are directed to production and use of chip arrays of polypeptides (including antibodies) and polynucleotides, and the disclosures of which are incorporated herein by reference in their entireties). By "solid support" is intended any solid support to which an antibody can be immobilized. Such solid supports include, but are not limited to nitrocellulose, diazocellulose, glass, polystyrene, polyvinylchloride, polycarbonate, polypropylene, polyethylene, dextran, Sepharose, agar, starch, nylon, beads and microtitre plates. Preferred are beads made of glass, latex or a magnetic material. Linkage of an antibody of the invention to a solid support can be accomplished by attaching one or both ends of the antibody to the support. Attachment may also be made at one or more internal sites in the antibody. Multiple attachments (both internal and at the ends of the antibody) may also be used according to the invention. Attachment can be via an amino acid linkage group such as a primary amino group, a carboxyl group, or a sulfhydryl (SH) group or by chemical linkage groups such as with cyanogen bromide (CNBr) linkage through a spacer. For non-covalent attachments, addition of an affinity tag sequence to the peptide can be used such as GST

(Smith, D.B., and Johnson, K.S., *Gene* 67:31 (1988)), polyhistidines (Hochuli, E., *et al.*, *J. Chromatog.* 411:77 (1987)), or biotin. Alternatively, attachment can be accomplished using a ligand which binds the Fc region of the antibodies of the invention, *e.g.*, protein A or protein G. Such affinity tags may be used for the reversible attachment of the antibodies to the support. Peptides may also be recognized via specific ligand-receptor interactions or using phage display methodologies that will be familiar to the skilled artisan, for their ability to bind polypeptides of the invention or fragments thereof.

Kits

In another aspect, the invention provides kits which may be used in producing the nucleic acid molecules, polypeptides, vectors, host cells, and antibodies, and in the recombinational cloning methods, of the invention. Kits according to this aspect of the invention may comprise one or more containers, which may contain one or more of the nucleic acid molecules, primers, polypeptides, vectors, host cells, or antibodies of the invention. In particular, a kit of the invention may comprise one or more components (or combinations thereof) selected from the group consisting of one or more recombination proteins (*e.g.*, Int) or auxiliary factors (*e.g.* IHF and/or Xis) or combinations thereof, one or more compositions comprising one or more recombination proteins or auxiliary factors or combinations thereof (for example, GATEWAY™ LR Clonase™ Enzyme Mix or GATEWAY™ BP Clonase™ Enzyme Mix) one or more Destination Vector molecules (including those described herein), one or more Entry Clone or Entry Vector molecules (including those described herein), one or more primer nucleic acid molecules (particularly those described herein), one or more host cells (*e.g.* competent cells, such as *E. coli* cells, yeast cells, animal cells (including mammalian cells, insect cells, nematode cells, avian cells, fish cells, etc.), plant cells, and most particularly *E. coli* DB3, DB3.1 (preferably *E. coli* LIBRARY EFFICIENCY® DB3.1™ Competent Cells; Life Technologies, Inc., Rockville, MD), DB4 and DB5; *see* U.S. Provisional Application No. 60/122,392, filed on March 2, 1999, and the corresponding U.S. Utility Application No. _____ of Hartley *et al.*, entitled "Cells Resistant to Toxic Genes and Uses Thereof," filed

on even day herewith, the disclosures of which are incorporated by reference herein in its entirety), and the like. In related aspects, the kits of the invention may comprise one or more nucleic acid molecules encoding one or more recombination sites or portions thereof, such as one or more nucleic acid molecules comprising a nucleotide sequence encoding the one or more recombination sites (or portions thereof) of the invention, and particularly one or more of the nucleic acid molecules contained in the deposited clones described herein. Kits according to this aspect of the invention may also comprise one or more isolated nucleic acid molecules of the invention, one or more vectors of the invention, one or more primer nucleic acid molecules of the invention, and/or one or more antibodies of the invention. The kits of the invention may further comprise one or more additional containers containing one or more additional components useful in combination with the nucleic acid molecules, polypeptides, vectors, host cells, or antibodies of the invention, such as one or more buffers, one or more detergents, one or more polypeptides having nucleic acid polymerase activity, one or more polypeptides having reverse transcriptase activity, one or more transfection reagents, one or more nucleotides, and the like. Such kits may be used in any process advantageously using the nucleic acid molecules, primers, vectors, host cells, polypeptides, antibodies and other compositions of the invention, for example in methods of synthesizing nucleic acid molecules (e.g., via amplification such as via PCR), in methods of cloning nucleic acid molecules (preferably via recombinational cloning as described herein), and the like.

Optimization of Recombinational Cloning System

The usefulness of a particular nucleic acid molecule, or vector comprising a nucleic acid molecule, of the invention in methods of recombinational cloning may be determined by any one of a number of assay methods. For example, Entry and Destination vectors of the present invention may be assessed for their ability to function (*i.e.*, to mediate the transfer of a nucleic acid molecule, DNA segment, gene, cDNA molecule or library from a cloning vector to an Expression Vector) by carrying out a recombinational cloning reaction as described in more detail in the Examples below and as described in U.S. Application Nos. 08/663,002, filed

June 7, 1996 (now U.S. Patent No. 5,888,732), 09/005,476, filed January 12, 1998, 09/177,387, filed October 23, 1998, and 60/108,324, filed November 13, 1998, the disclosures of which are incorporated by reference herein in their entirety. Alternatively, the functionality of Entry and Destination Vectors prepared according to the invention may be assessed by examining the ability of these vectors to recombine and create cointegrate molecules, or to transfer a nucleic acid molecule of interest, using an assay such as that described in detail below in Example 19. Analogously, the formulation of compositions comprising one or more recombination proteins or combinations thereof, for example GATEWAY™ LR Clonase™ Enzyme Mix and GATEWAY™ BP Clonase™ Enzyme Mix, may be optimized using assays such as those described below in Example 18.

Uses

There are a number of applications for the compositions, methods and kits of the present invention. These uses include, but are not limited to, changing vectors, targeting gene products to intracellular locations, cleaving fusion tags from desired proteins, operably linking nucleic acid molecules of interest to regulatory genetic sequences (*e.g.*, promoters, enhancers, and the like), constructing genes for fusion proteins, changing copy number, changing replicons, cloning into phages, and cloning, *e.g.*, PCR products, genomic DNAs, and cDNAs. In addition, the nucleic acid molecules, vectors, and host cells of the invention may be used in the production of polypeptides encoded by the nucleic acid molecules, in the production of antibodies directed against such polypeptides, in recombinational cloning of desired nucleic acid sequences, and in other applications that may be enhanced or facilitated by the use of the nucleic acid molecules, vectors, and host cells of the invention.

In particular, the nucleic acid molecules, vectors, host cells, polypeptides, antibodies, and kits of the invention may be used in methods of transferring one or more desired nucleic acid molecules or DNA segments, for example one or more genes, cDNA molecules or cDNA libraries, into a cloning or Expression Vector for use in transforming additional host cells for use in cloning or

amplification of, or expression of the polypeptide encoded by, the desired nucleic acid molecule or DNA segment. Such recombinational cloning methods which may advantageously use the nucleic acid molecules, vectors, and host cells of the invention, are described in detail in the Examples below, and in commonly owned
5 U.S. Application Nos. 08/486,139, filed June 7, 1995, 08/663,002, filed June 7, 1996 (now U.S. Patent No. 5,888,732), 09/005,476, filed January 12, 1998, 09/177,387, filed October 23, 1998, and 60/108,324, filed November 13, 1998, the disclosures of all of which are incorporated by reference herein in their entireties.

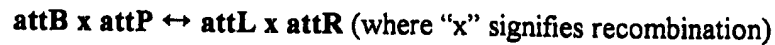
10 It will be understood by one of ordinary skill in the relevant arts that other suitable modifications and adaptations to the methods and applications described herein are readily apparent from the description of the invention contained herein in view of information known to the ordinarily skilled artisan, and may be made
15 without departing from the scope of the invention or any embodiment thereof. Having now described the present invention in detail, the same will be more clearly understood by reference to the following examples, which are included herewith for purposes of illustration only and are not intended to be limiting of the invention.

20 *Examples*

Example 1: Recombination Reactions of Bacteriophage λ

25 The *E. coli* bacteriophage λ can grow as a lytic phage, in which case the host cell is lysed, with the release of progeny virus. Alternatively, lambda can integrate into the genome of its host by a process called lysogenization (see Figure 60). In this lysogenic state, the phage genome can be transmitted to daughter cells for many generations, until conditions arise that trigger its excision from the genome.
30 At this point, the virus enters the lytic part of its life cycle. The control of the switch between the lytic and lysogenic pathways is one of the best understood processes in molecular biology (M. Ptashne, *A Genetic Switch*, Cell Press, 1992).

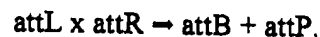
The integrative and excisive recombination reactions of λ , performed *in vitro*, are the basis of Recombinational Cloning System of the present invention. They can be represented schematically as follows:



10 The four att sites contain binding sites for the proteins that mediate the reactions. The wild type attP, attB, attL, and attR sites contain about 243, 25, 100, and 168 base pairs, respectively. The attB x attP reaction (hereinafter referred to as a "BP Reaction," or alternatively and equivalently as an "Entry Reaction" or a "Gateward Reaction") is mediated by the proteins Int and IHF. The attL x attR reaction (hereinafter referred to as an "LR Reaction," or alternatively and equivalently as a "Destination Reaction") is mediated by the proteins Int, IHF, and Xis. Int (integrase) and Xis (excisionase) are encoded by
15 the λ genome, while IHF (integration host factor) is an *E. coli* protein. For a general review of lambda recombination, see: A. Landy, *Ann. Rev. Biochem.* 58: 913-949 (1989).

20 ***Example 2: Recombination Reactions of the Recombinational Cloning System***

 The LR Reaction -- the exchange of a DNA segment from an Entry Clone to a Destination Vector -- is the *in vitro* version of the λ excision reaction:



30 There is a practical imperative for this configuration: after an LR Reaction in one configuration of the present method, an att site usually separates a functional motif (such as a promoter or a fusion tag) from a nucleic acid molecule of interest in an Expression Clone, and the 25 bp attB site is much smaller than the attP, attL, and attR sites.

 Note that the recombination reaction is conservative, i.e., there is no net synthesis or loss of base pairs. The DNA segments that flank the recombination

sites are merely switched. The wild type λ recombination sites are modified for purposes of the GATEWAY™ Cloning System, as follows:

To create certain preferred Destination Vectors, a part (43 bp) of attR was removed, to make the excise reaction irreversible and more efficient (W. Bushman et al., *Science* 230: 906, 1985). The attR sites in preferred Destination Vectors of the invention are 125 bp in length. Mutations were made to the core regions of the att sites, for two reasons: (1) to eliminate stop codons, and (2) to ensure specificity of the recombination reactions (i.e., attR1 reacts only with attL1, attR2 reacts only with attL2, etc.).

Other mutations were introduced into the short (5 bp) regions flanking the 15 bp core regions of the attB sites to minimize secondary structure formation in single-stranded forms of attB plasmids, e.g., in phagemid ssDNA or in mRNA. Sequences of attB1 and attB2 to the left and right of a nucleic acid molecule of interest after it has been cloned into a Destination Vector are given in Figure 6.

Figure 61 illustrates how an Entry Clone and a Destination Vector recombine in the LR Reaction to form a co-integrate, which resolves through a second reaction into two daughter molecules. The two daughter molecules have the same general structure regardless of which pair of sites, attL1 and attR1 or attL2 and attR2, react first to form the co-integrate. The segments change partners by these reactions, regardless of whether the parental molecules are both circular, one is circular and one is linear, or both are linear. In this example, selection for ampicillin resistance carried on the Destination Vector, which also carries the death gene *ccdB*, provides the means for selecting only for the desired attB product plasmid.

Example 3: Protein Expression in the Recombinational Cloning System

Proteins are expressed *in vivo* as a result of two processes, transcription (DNA into RNA), and translation (RNA into protein). For a review of protein expression in prokaryotes and eukaryotes, see Example 13 below. Many vectors (pUC, BlueScript, pGem) use interruption of a transcribed *lacZ* gene for blue-white screening. These plasmids, and many Expression Vectors, use the *lac* promoter to control expression of cloned genes. Transcription from the *lac*

promoter is turned on by adding the inducer IPTG. However, a low level of RNA is made in the absence of inducer, i.e., the lac promoter is never completely off. The result of this "leakiness" is that genes whose expression is harmful to *E. coli* may prove difficult or impossible to clone in vectors that contain the lac promoter, or they may be cloned only as inactive mutants.

In contrast to other gene expression systems, nucleic acid molecules cloned into an Entry Vector may be designed *not* to be expressed. The presence of the strong transcriptional terminator *rrnB* (Orosz, et al., *Eur. J. Biochem.* 201: 653, 1991) just upstream of the attL1 site keeps transcription from the vector promoters (drug resistance and replication origin) from reaching the cloned gene. However, if a toxic gene is cloned into a Destination Vector, the host may be sick, just as in other expression systems. But the reliability of subcloning by *in vitro* recombination makes it easier to recognize that this has happened -- and easier to try another expression option in accordance with the methods of the invention, if necessary.

Example 4: Choosing the Right Entry Vector

There are two kinds of choices that must be made in choosing the best Entry Vector, dictated by (1) the particular DNA segment that is to be cloned, and (2) what is to be accomplished with the cloned DNA segment. These factors are critical in the choice of Entry Vector used, because when the desired nucleic acid molecule of interest is moved from the Entry Vector to a Destination Vector, all the base pairs between the nucleic acid molecule of interest and the Int cutting sites in attL1 and attL2 (such as in Figure 6) move into the Destination Vector as well. For genomic DNAs that are not expressed as a result of moving into a Destination Vector, these decisions are not as critical.

For example, if an Entry Vector with certain translation start signals is used, those sequences will be translated into amino acids if an amino-terminal fusion to the desired nucleic acid molecule of interest is made. Whether the desired nucleic acid molecule of interest is to be expressed as fusion protein, native protein, or both, dictates whether translational start sequences must be included between the attB sites of the clone (native protein) or, alternatively, supplied by the Destination

Vector (fusion protein). In particular, Entry Clones that include translational start sequences may prove less suitable for making fusion proteins, as internal initiation of translation at these sites can decrease the yield of N-terminal fusion protein. These two types of expression afforded by the compositions and methods of the invention are illustrated in Figure 62.

No Entry Vector is likely to be optimal for all applications. The nucleic acid molecule of interest may be cloned into any of several optimal Entry Vectors.

As an example, consider pENTR7 (Figure 16) and pENTR11 (Figure 20), which are useful in a variety of applications, including (but not limited to):

- Cloning cDNAs from most of the commercially available libraries. The sites to the left and right of the *ccdB* death gene have been chosen so that directional cloning is possible if the DNA to be cloned does not have two or more of these restriction sites.

- Cloning of genes directionally: *Sal*I, *Bam*HI, *Xmn*I (blunt), or *Kpn*I on the left of *ccdB*; *Not*I, *Xho*I, *Xba*I, or *Eco*RV (blunt), on the right.

- Cloning of genes or gene fragments with a blunt amino end at the *Xmn*I site. The *Xmn*I site has four of the six most favored bases for eukaryotic expression (see Example 13, below), so that if the first three bases of the DNA to be cloned are ATG, the open reading frame (ORF) will be expressed in eukaryotic cells (e.g., mammalian cells, insect cells, yeast cells) when it is transcribed in the appropriate Destination Vector. In addition, in pENTR11, a Shine-Dalgarno sequence is situated 8 bp upstream, for initiating protein synthesis in a prokaryotic host cell (particularly a bacterial cell, such as *E. coli*) at an ATG.

- Cleaving off amino terminal fusions (e.g., His₆, GST, or thioredoxin) using the highly specific TEV (Tobacco Etch Virus) protease (available from Life Technologies, Inc.). If the nucleic acid molecule of interest is cloned at the

blunt *Xmn*I site, TEV cleavage will leave two amino acids on the amino end of the expressed protein.

•Selecting against uncut or singly cut Entry Vector molecules during cloning with restriction enzymes and ligase. If the *ccdB* gene is not removed with a double digest, it will kill any recipient *E. coli* cell that does not contain a mutation that makes the cell resistant to *ccdB* (see U.S. Provisional Application No. 60/122,392, filed on March 2, 1999, the disclosure of which is incorporated by reference herein in its entirety).

•Allowing production of amino fusions with ORFs in all cloning sites. There are no stop codons (in the attL1 reading frame) upstream of the *ccdB* gene.

In addition, pENTR11 is also useful in the following applications:

•Cloning cDNAs that have an *Nco*I site at the initiating ATG into the *Nco*I site. Similar to the *Xmn*I site, this site has four of the six most favored bases for eukaryotic expression. Also, a Shine-Dalgarno sequence is situated 8 bp upstream, for initiating protein synthesis in a prokaryotic host cell (particularly a bacterial cell, such as *E. coli*) at an ATG.

•Producing carboxy fusion proteins with ORFs positioned in phase with the reading frame convention for carboxy-terminal fusions (see Figure 20A).

Table 1 lists some non-limiting examples of Entry Vectors and their characteristics, and Figures 10-20 show their cloning sites. All of the Entry Vectors listed in Table 1 are available commercially from Life Technologies, Inc., Rockville, Maryland. Other Entry Vectors not specifically listed here, which comprise alternative or additional features may be made by one of ordinary skill using routine methods of molecular and cellular biology, in view of the disclosure contained herein.

Table 1 Examples of Entry Vectors

Designation	Mnemonic Name	Class of Entry Vector	Distinctive Cloning Sites	Amino Fusions	Native Protein in E.coli	Native Protein in Eukaryotic Cells	Protein Synthesis Features
pENTR-1A, 2B, 3C	Minimal blunt RF A, B, C	Alternative Reading Frame Vectors	Reading frame A, B, or C; blunt cut closest to attL1	Good	Poor	Good	Minimal amino acids between tag and protein; no SD
pENTR4	Minimal Nco	Restr. Enz. Cleavage Vectors	Nco I site (common in euk. cDNAs) closest to attL1	Good	Poor	Good	Good Kozac; no SD
pENTR5	Minimal Nde	Restr. Enz. Cleavage Vectors	Nde I site closest to attL1	Good	Poor	Poor at Nde I, Good at Xmn I	No SD; poor Kozac at Nde, good at Xmn
pENTR6	Minimal Sph	Restr. Enz. Cleavage Vectors	Sph I site closest to attL1	Good	Poor	Poor at Sph I, Good at Xmn I	No SD; poor Kozac at Sph, good at Xmn
pENTR7	TEV Blunt	TEV Cleavage Site Present	Xmn I (blunt) is first cloning site after TEV site	Good	Poor	Good at Xmn I site	TEV protease leaves Gly-Thr on amino end of protein; no SD
pENTR8	TEV Nco	TEV Cleavage Site Present	Nco I is first cloning site after TEV site	Good	Poor	Good	TEV protease leaves Gly-Thr on amino end of protein; no SD

pENTR9	TEV Nde	TEV Cleavage Site Present	Nde I is first cloning site after TEV site	Good	Poor	Poor	TEV protease leaves Gly-Thr on amino end of protein; no SD, poor Kozac
pENTR10	Nde with SD	Good SD for E.coli Expression	Strong SD; Nde I site, no TEV	Poor	Good	Poor	Strong SD, internal starts in amino fusions. Poor Kz. No TEV
pENTR11	2 X SD+Kozac	Good SD for E.coli Expression	Xmn I (blunt) and Nco I sites each preceded by SD and Kozac	Good	Good	Good	Strong SD/Koz Internal starts in amino fusions. No TEV

Entry vectors pENTR1A (Figures 10A and 10B), pENTR2B (Figures 11A and 11B), and pENTR3C (Figures 12A and 12B) are almost identical, except that the restriction sites are in different reading frames. Entry vectors pENTR4 (Figures 13A and 13B), pENTR5 (Figures 14A and 14B), and pENTR6 (Figures 15A and 15B) are essentially identical to pENTR1A, except that the blunt *DraI* site has been replaced with sites containing the ATG methionine codon: *NcoI* in pENTR4, *NdeI* in pENTR5, and *SphI* in pENTR6. Nucleic acid molecules that contain one of these sites at the initiating ATG can be conveniently cloned in these Entry vectors. The *NcoI* site in pENTR4 is especially useful for expression of nucleic acid molecules in eukaryotic cells, since it contains many of the bases that give efficient translation (see Example 13, below). (Nucleic acid molecules of interest cloned into the *NdeI* site of pENTR5 are not expected to be highly expressed in eukaryotic cells, because the cytosine at position -3 from the initiating ATG is rare in eukaryotic genes.)

Entry vectors pENTR7 (Figures 16A and 16B), pENTR8 (Figures 17A and 17B), and pENTR9 (Figures 18A and 18B) contain the recognition site for the TEV protease between the attL1 site and the cloning sites. Cleavage sites for *XmnI* (blunt), *NcoI*, and *NdeI*, respectively, are the most 5' sites in these Entry vectors. Amino fusions can be removed efficiently if nucleic acid molecules are cloned into these Entry vectors. TEV protease is highly active and highly specific.

Example 5: Controlling Reading Frame

One of the trickiest tasks in expression of cloned nucleic acid molecules is making sure the reading frame is correct. (Reading frame is important if fusions are being made between two ORFs, for example between a nucleic acid molecule of interest and a His6 or GST domain.) For purposes of the present invention, the following convention has been adopted: The reading frame of the DNA cloned into any Entry Vector must be in phase with that of the attB1 site shown in Figure 16A, pENTR7. Notice that the six As of the attL1 site are split into two lysine codons (aaa aaa). The Destination Vectors that make amino fusions were constructed such that they enter the attR1 site in this reading frame.

Destination Vectors for carboxy terminal fusions were also constructed, including those containing His₆ (pDEST23; Figure 43), GST (pDEST24; Figure 44), or thioredoxin (pDEST25; Figure 45) C-terminal fusion sequences.

Therefore, if a nucleic acid molecule of interest is cloned into an Entry Vector so that the aaa aaa reading frame within the attL1 site is in phase with the nucleic acid molecule's ORF, amino terminal fusions will automatically be correctly phased, for all the fusion tags. This is a significant improvement over the usual case, where each different vector can have different restriction sites and different reading frames.

See Example 15 for a practical example of how to choose the most appropriate combinations of Entry Vector and Destination Vector.

Materials

Unless otherwise indicated, the following materials were used in the remaining Examples included herein:

5X LR Reaction Buffer:

200-250 mM (preferably 250 mM) Tris-HCl, pH 7.5

250-350 mM (preferably 320 mM) NaCl

1.25-5 mM (preferably 4.75 mM) EDTA

12.5-35 mM (preferably 22-35 mM, and most preferably 35 mM)

Spermidine-HCl

1 mg/ml bovine serum albumin

GATEWAY™ LR Clonase™ Enzyme Mix:

per 4 µl of 1X LR Reaction Buffer:

150 ng carboxy-His6-tagged Int (see U.S. Appl. Nos. 60/108,324, filed

November 13, 1998, and 09/438,358, filed November 12,

1999, both entirely incorporated by reference herein)

-106-

25 ng carboxy-His6-tagged Xis (see U.S. Appl. Nos. 60/108,324, filed
November 13, 1998, and 09/438,358, filed November 12,
1999, both entirely incorporated by reference herein)

30 ng IHF

50% glycerol

5X BP Reaction Buffer:

125 mM Tris-HCl, pH 7.5

110 mM NaCl

25 mM EDTA

25 mM Spermidine-HCl

5 mg/ml bovine serum albumin

GATEWAY™ BP Clonase™ Enzyme Mix:

per 4 µl of 1X BP Reaction Buffer:

200 ng carboxy-His6-tagged Int (see U.S. Appl. Nos. 60/108,324, filed
November 13, 1998, and 09/438,358, filed November 12,
1999, both entirely incorporated by reference herein)

80 ng IHF

50% glycerol

10X Clonase Stop Solution:

50 mM Tris-HCl, pH 8.0

1 mM EDTA

2 mg/ml Proteinase K

Example 6: LR ("Destination") Reaction

To create a new Expression Clone containing the nucleic acid molecule of
interest (and which may be introduced into a host cell, ultimately for production
of the polypeptide encoded by the nucleic acid molecule), an Entry Clone or
Vector containing the nucleic acid molecule of interest, prepared as described

herein, is reacted with a Destination Vector. In the present example, a β -Gal gene flanked by attL sites is transferred from an Entry Clone to a Destination Vector.

Materials needed:

- 5 X LR Reaction buffer
- Destination Vector (preferably linearized), 75-150 ng/ μ l
- Entry Clone containing nucleic acid molecule of interest, 100-300 ng in $\leq 8 \mu$ l TE buffer
- Positive control Entry Clone (pENTR- β -Gal) DNA (See note, below)
- Positive control Destination Vector, pDEST1 (pTrc), 75 ng/ μ l
- GATEWAY™ LR Clonase™ Enzyme Mix (stored at - 80° C)
- 10X Clonase Stop solution
- pUC19 DNA, 10 pg/ μ l
- Chemically competent *E. coli* cells (competence: $\geq 1 \times 10^7$ CFU/ μ g), 400 μ l.
- LB Plates containing ampicillin (100 μ g/ml) and methicillin (200 μ g/ml) \pm X-gal and IPTG (See below)

Notes:

Preparation of the Entry Clone DNA: Miniprep DNA that has been treated with RNase works well. A reasonably accurate quantitation ($\pm 50\%$) of the DNA to be cloned is advised, as the GATEWAY™ reaction appears to have an optimum of about 100-300 ng of Entry Clone per 20 μ l of reaction mix.

The positive control Entry Clone, pENTR- β -Gal, permits functional analysis of clones based on the numbers of expected blue vs. white colonies on LB plates containing IPTG + Blueo-gal (or X-gal), in addition to ampicillin (100 μ g/ml) and methicillin (200 μ g/ml). Because β -Galactosidase is a large protein, it often yields a less prominent band than many smaller proteins do on SDS protein gels.

In the Positive Control Entry Vector pENTR- β -Gal, the coding sequence of β -Gal has been cloned into pENTR11 (Figures 20A and 20B), with translational start signals permitting expression in *E. coli*, as well as in eukaryotic

cells. The positive control Destination Vector, for example pDEST1 (Figure 21), is preferably linearized.

To prepare X-gal + IPTG plates, either of the following protocols may be used:

5

A. With a glass rod, spread over the surface of an LB agar plate: 40 μ l of 20 mg/ml X-gal (or Blueo-gal) in DMF plus 4 μ l 200 mg/ml IPTG. Allow liquid to adsorb into agar for 3-4 hours at 37° C before plating cells.

10

B. To liquid LB agar at ~45°C, add: X-gal (or Blueo-Gal) (20 mg/ml in DMF) to make 50 μ g/ml and IPTG (200 mM in water) to make 0.5-1 mM, just prior to pouring plates. Store X-gal and Blueo-Gal in a light-shielded container.

15

Colony color may be enhanced by placing the plates at 5°C for a few hours after the overnight incubation at 37°C. Protocol B can give more consistent colony color than A, but A is more convenient when selection plates are needed on short notice.

20

Recombination in Clonase reactions continues for many hours. While incubations of 45-60 minutes are usually sufficient, reactions with large DNAs, or in which both parental DNAs are supercoiled, or which will be transformed into cells of low competence, can be improved with longer incubation times, such as 2-24 hours at 25°C.

Procedure:

25

1. Assemble reactions as follows (combine all components at room temperature, except GATEWAY™ LR Clonase™ Enzyme Mix ("Clonase LR"), before removing Clonase LR from frozen storage):

Component	Tube 1	Tube 2	Tube 3	Tube 4
	Neg.	Pos.	Neg.	Test
p-Gate-βGal, (Positive control Entry Clone) 75 ng/μl	4 μl	4 μl		
pDEST1 (Positive control Destination Vector), 75 ng/μl	4 μl	4 μl		
Your Entry Clone (100-300 ng)			1 - 8 μl	1 - 8 μl
Destination Vector for your nucleic acid molecule, 75 ng/μl			4 μl	4 μl
5 X LR Reaction Buffer	4 μl	4 μl	4 μl	4 μl
TE	8 μl	4 μl	To 20 μl	To 16 μl
GATEWAY™ LR Clonase™ Enzyme Mix (store at - 80° C, add last)	---	4 μl	---	4 μl
Total Volume	20 μl	20 μl	20 μl	20 μl

2. Remove the GATEWAY™ LR Clonase™ Enzyme Mix from the -80° C freezer, place immediately on ice. The Clonase takes only a few minutes to thaw.
3. Add 4 μl of GATEWAY™ LR Clonase™ Enzyme Mix to reactions #2 and #4;
4. Return GATEWAY™ LR Clonase™ Enzyme Mix to - 80° C freezer.
5. Incubate tubes at 25° for at least 60 minutes.
6. Add 2 μl Clonase Stop solution to all reactions. Incubate for 20 min at 37° C. (This step usually increases the total number of colonies obtained by 10-20 fold.)
7. Transform 2 μl into 100 μl competent *E. coli*. Select on plates containing ampicillin at 100 μg/ml.

Example 7: Transformation of *E. coli*

To introduce cloning or Expression Vectors prepared using the recombinational cloning system of the invention, any standard *E. coli* transformation protocol should be satisfactory. The following steps are recommended for best results:

-110-

1. Let the mixture of competent cells and Recombinational Cloning System reaction product stand on ice at least 15 minutes prior to the heat-shock step. This gives time for the recombination proteins to dissociate from the DNA, and improves the transformation efficiency.

2. Expect the reaction to be about 1%-5% efficient, i.e., 2 μ l of the reaction should contain at least 100 pg of the Expression Clone plasmid (taking into account the amounts of each parental plasmid in the reaction, and the subsequent dilution). If the E. coli cells have a competence of 10^7 CFU/ μ g, 100 pg of the desired clone plasmid will give about 1000 colonies, or more, if the entire transformation is spread on one ampicillin plate.

3. Always do a control pUC DNA transformation. If the number of colonies is not what you expect, the pUC DNA transformation gives you an indication of where the problem was.

Example 8: Preparation of attB-PCR Product

For preparation of attB-PCR products in the PCR cloning methods described in Example 9 below, PCR primers containing attB1 and attB2 sequences are used. The attB1 and attB2 primer sequences are as follows:

attB1: 5'-GGGGACAAGTTTGTACAAAAAAGCAGGCT-(template-specific sequence)-3'

attB2: 5'-GGGGACCACTTTGTACAAGAAAGCTGGGT-(template-specific sequence)-3'

The attB1 sequence should be added to the amino primer, and the attB2 sequence to the carboxy primer. The 4 guanines at the 5' ends of each of these primers enhance the efficiency of the minimal 25 bp attB sequences as substrates for use in the cloning methods of the invention.

Standard PCR conditions may be used to prepare the PCR product. The following suggested protocol employs PLATINUM *Taq* DNA Polymerase High

-111-

Fidelity®, available commercially from Life Technologies, Inc. (Rockville, MD). This enzyme mix eliminates the need for hot starts, has improved fidelity over Taq, and permits synthesis of a wide range of amplicon sizes, from 200 bp to 10 kb, or more, even on genomic templates.

Materials needed:

- PLATINUM Taq DNA Polymerase High Fidelity® (Life Technologies, Inc.)
- attB1- and attB2- containing primer pair (see above) specific for your template
- DNA template (linearized plasmid or genomic DNA)
- 10X High Fidelity PCR Buffer
- 10 mM dNTP mix
- PEG/MgCl₂ Mix (30% PEG 8000, 30 mM MgCl₂)

Procedure:

1.) Assemble the reaction as follows:

Component	Reaction with <u>Plasmid Target</u>	Reaction with <u>Genomic Target</u>
10X High Fidelity PCR Buffer	5 µl	5 µl
dNTP Mix 10 mM	1 µl	1 µl
MgSO ₄ , 50mM	2 µl	2 µl
attB1 Primer, 10 µM	2 µl	1 µl
attB2 Primer, 10 µM	2 µl	1 µl
Template DNA	1-5 ng*	≥ 100 ng
PLATINUM Taq High Fidelity	2 µl	1 µl
Water	to 50 µl	to 50 µl

* Use of higher amounts of plasmid template may permit fewer cycles (10-15) of PCR

2.) Add 2 drops mineral oil, as appropriate.

3.) Denature for 30 sec. at 94°C.

4.) Perform 25 cycles:

94°C for 15 sec-30 sec

55°C for 15 sec-30 sec

68°C for 1 min per kb of template.

5.) Following the PCR reaction, apply 1-2 µl of the reaction mixture to an agarose gel, together with size standards (*e.g.*, 1 Kb Plus Ladder, Life Technologies, Inc.) and quantitation standards (*e.g.*, Low Mass Ladder, Life Technologies, Inc.), to assess the yield and uniformity of the product.

Purification of the PCR product is recommended, to remove attB primer dimers which can clone efficiently into the Entry Vector. The following protocol is fast and will remove DNA <300 bp in size:

6.) Dilute the 50 µl PCR reaction to 200 µl with TE.

7.) Add 100 µl PEG/MgCl₂ Solution. Mix and centrifuge immediately at 13,000 RPM for 10 min at room temperature. Remove the supernatant (pellet is clear and hard to see).

8.) Dissolve the pellet in 50 µl TE and check recovery on a gel.

If the starting PCR template is a plasmid that contains the gene for Kan^r, it is advisable to treat the completed PCR reaction with the restriction enzyme *DpnI*, to degrade the plasmid since unreacted residual starting plasmid is a potential source of false-positive colonies from the transformation of the GATEWAY™ Cloning System reaction. Adding ~5 units of *DpnI* to the completed PCR reaction and incubating for 15 min at 37°C will eliminate this potential problem. Heat inactivate the *DpnI* at 65°C for 15 min, prior to using the PCR product in the GATEWAY™ Cloning System reaction.

Example 9: Cloning attB-PCR products into Entry Vectors via the BP ("Gateway") Reaction

The addition of 5'-terminal attB sequences to PCR primers allows synthesis of a PCR product that is an efficient substrate for recombination with a Donor (attP) Plasmid in the presence of GATEWAY™ BP Clonase™ Enzyme Mix. This reaction produces an Entry Clone of the PCR product (See Figure 8).

The conditions of the Gateway Cloning reaction with an attB PCR substrate are similar to those of the BP Reaction (see Example 10 below), except that the attB-PCR product (see Example 8) substitutes for the Expression Clone, and the attB-PCR positive control (attB-tet^r) substitutes for the Expression Clone Positive Control (GFP).

Materials needed:

- 5 X BP Reaction Buffer
- Desired attB-PCR product DNA, 50-100 ng in $\leq 8 \mu\text{l}$ TE.
- Donor (attP) Plasmid (Figures 49-54), 75 ng/ μl , supercoiled DNA
- attB-tet^r PCR product positive control, 25 ng/ μl
- GATEWAY™ BP Clonase™ Enzyme Mix (stored at - 80° C)
- 10x Clonase Stop Solution
- pUC19 DNA, 10 pg/ μl .
- Chemically competent E.coli cells (competence: $\geq 1 \times 10^7$ CFU/ μg), 400 μl

Notes:

- Preparation of attB-PCR DNA: see Example 8.
- The Positive Control attB-tet^r PCR product contains a functional copy of the tet^r gene of pBR322, with its own promoter. By plating the transformation of the control BP Reaction on kanamycin (50 $\mu\text{g/ml}$) plates (if kan^r Donor Plasmids are used; see Figures 49-52) or an alternative selection agent (e.g., gentamycin, if gen^r Donor Plasmids are used; see Figure 54), and then picking about 50 of these colonies onto plates with tetracycline (20 $\mu\text{g/ml}$), the

percentage of Entry Clones containing functional tet^r among the colonies from the positive control reaction can be determined (% Expression Clones = (number of tet^r + kan^r (or gen^r) colonies/kan^r (or gen^r) colonies).

Procedure:

1. Assemble reactions as follows. Combine all components except GATEWAY™ BP Clonase™ Enzyme Mix, before removing GATEWAY™ BP Clonase™ Enzyme Mix from frozen storage.

	Neg.	Pos.	Test
Component	Tube 1	Tube 2	Tube 3
attB-PCR product, 50-100 ng			1 - 8 µl
Donor (attP) Plasmid 75 ng/µl	2 µl	2 µl	2 µl
attB-PCR tet ^r control DNA (75 ng/µl)		4 µl	
5 X BP Reaction Buffer	4 µl	4 µl	4 µl
TE	10 µl	6 µl	To 16 µl
GATEWAY™ BP Clonase™ Enzyme Mix (store at -80° C, add last)	4 µl	4 µl	4 µl
Total Volume	20 µl	20 µl	20 µl

2. Remove the GATEWAY™ BP Clonase™ Enzyme Mix from the -80° C freezer, place immediately on ice. The Clonase takes only a few minutes to thaw.
3. Add 4 µl of GATEWAY™ BP Clonase™ Enzyme Mix to the subcloning reaction, mix.
4. Return GATEWAY™ BP Clonase™ Enzyme Mix to - 80° C freezer.
5. Incubate tubes at 25° for at least 60 minutes.

6. Add 2 μ l Proteinase K (2 μ g/ μ l) to all reactions. Incubate for 20 min at 37°C.
7. Transform 2 μ l into 100 μ l competent *E. coli*, as per 3.2, above. Select on LB plates containing kanamycin, 50 μ g/ml.

Results:

In initial experiments, primers for amplifying tetR and ampR from pBR322 were constructed containing only the tetR- or ampR-specific targeting sequences, the targeting sequences plus attB1 (for forward primers) or attB2 (for reverse primers) sequences shown in Figure 9, or the attB1 or attB2 sequences with a 5' tail of four guanines. The construction of these primers is depicted in Figure 65. After PCR amplification of tetR and ampR from pBR322 using these primers and cloning the PCR products into host cells using the recombinational cloning system of the invention, the results shown in Figure 66 were obtained. These results demonstrated that primers containing attB sequences provided for a somewhat higher number of colonies on the tetracycline and ampicillin plates. However, inclusion of the 5' extensions of four or five guanines on the primers in addition to the attB sequences provided significantly better cloning results, as shown in Figures 66 and 67. These results indicate that the optimal primers for cloning of PCR products using recombinational cloning will contain the recombination site sequences with a 5' extension of four or five guanine bases.

To determine the optimal stoichiometry between attB-containing PCR products and attP-containing Donor plasmid, experiments were conducted where the amount of PCR product and Donor plasmid were varied during the BP Reaction. Reaction mixtures were then transformed into host cells and plated on tetracycline plates as above. Results are shown in Figure 68. These results indicate that, for optimal recombinational cloning results with a PCR product in the size range of the tet gene, the amounts of attP-containing Donor plasmids are between about 100-500 ng (most preferably about 200-300 ng), while the optimal concentrations of attB-containing PCR products is about 25-100 ng (most preferably about 100 ng), per 20 μ l reaction.

Experiments were then conducted to examine the effect of PCR product size on efficiency of cloning via the recombinational cloning approach of the invention.

PCR products containing attB1 and attB2 sites, at sizes 256 bp, 1 kb, 1.4 kb, 3.4 kb, 4.6 kb, 6.9 kb and 10.1 kb were prepared and cloned into Entry vectors as described above, and host cells were transformed with the Entry vectors containing the cloned PCR products. For each PCR product, cloning efficiency was calculated relative to cloning of pUC19 positive control plasmids as follows:

$$\text{Cloning Efficiency} = \frac{\text{CFU/ng attB PCR product}}{\text{CFU/ng pUC19 control}} \times \frac{\text{Size (kb) PCR product}}{\text{Size (kb) pUC19 control}}$$

The results of these experiments are depicted in Figures 69A-69C (for 256 bp PCR fragments), 70A-70C (for 1 kb PCR fragments), 71A-71C (for 1.4 kb PCR fragments), 72A-72C (for 3.4 kb PCR fragments), 73A-73C (for 4.6 kb PCR fragments), 74 (for 6.9 kb PCR fragments), and 75-76 (for 10.1 kb PCR fragments). The results shown in these figures are summarized in Figure 77, for different weights and moles of input PCR DNA.

Together, these results demonstrate that attB-containing PCR products ranging in size from about 0.25 kb to about 5 kb clone relatively efficiently in the recombinational cloning system of the invention. While PCR products larger than about 5 kb clone less efficiently (apparently due to slow resolution of cointegrates), longer incubation times during the recombination reaction appears to improve the efficiency of cloning of these larger PCR fragments. Alternatively, it may also be possible to improve efficiency of cloning of large (> about 5 kb) PCR fragments by using lower levels of input attP Donor plasmid and perhaps attB-containing PCR product, and/or by adjusting reaction conditions (e.g., buffer conditions) to favor more rapid resolution of the cointegrates.

Example 10: The BP Reaction

One purpose of the Gateward ("Entry") reaction is to convert an Expression Clone into an Entry Clone. This is useful when you have isolated an individual Expression Clone from an Expression Clone cDNA library, and you wish to transfer the nucleic acid molecule of interest into another Expression Vector, or

to move a population of molecules from an attB or attL library. Alternatively, you may have mutated an Expression Clone and now wish to transfer the mutated nucleic acid molecule of interest into one or more new Expression Vectors. In both cases, it is necessary first to convert the nucleic acid molecule of interest to an Entry Clone.

Materials needed:

- 5 X BP Reaction Buffer
- Expression Clone DNA, 100-300 ng in $\leq 8 \mu\text{l}$ TE.
- Donor (attP) Vector, 75 ng/ μl , supercoiled DNA
- Positive control attB-tet-PCR DNA, 25 ng/ μl
- GATEWAY™ BP Clonase™ Enzyme Mix (stored at -80°C)
- Clonase Stop Solution (Proteinase K, 2 $\mu\text{g}/\mu\text{l}$).

Notes:

Preparation of the Expression Clone DNA: Miniprep DNA treated with RNase works well.

1. As with the LR Reaction (see Example 14), the BP Reaction is strongly influenced by the topology of the reacting DNAs. In general, the reaction is most efficient when one of the DNAs is linear and the other is supercoiled, compared to reactions where the DNAs are both linear or both supercoiled. Further, linearizing the attB Expression Clone (anywhere within the vector) will usually give more colonies than linearizing the Donor (attP) Plasmid. If finding a suitable cleavage site within your Expression Clone vector proves difficult, you may linearize the Donor (attP) Plasmid between the attP1 and attP2 sites (for example, at the *Nco*I site), avoiding the *ccdB* gene. Maps of Donor (attP) Plasmids are given in Figures 49-54.

Procedure:

1. Assemble reactions as follows. Combine all components at room temperature, except GATEWAY™ BP Clonase™ Enzyme Mix, before removing GATEWAY™ BP Clonase™ Enzyme Mix from freezer.

	Neg.	Pos.	Test
Component	Tube 1	Tube 2	Tube 3
Positive Control, attB-tet-PCR DNA, 25 ng/ μ l	4 μ l	4 μ l	
Desired attB Expression Clone DNA (100ng) linearized			1 - 8 μ l
Donor (attP) Plasmid, 75 ng/ μ l	2 μ l	2 μ l	2 μ l
5 X BP Reaction Buffer	4 μ l	4 μ l	4 μ l
TE	10 μ l	6 μ l	To 16 μ l
GATEWAY™ BP Clonase™ Enzyme Mix (store at - 80° C, add last)	---	4 μ l	4 μ l
Total Volume	20 μ l	20 μ l	20 μ l

2. Remove the GATEWAY™ BP Clonase™ Enzyme Mix from the -80°C freezer, place immediately on ice. The mixture takes only a few minutes to thaw.
3. Add 4 μ l of GATEWAY™ BP Clonase™ Enzyme Mix to the subcloning reaction, mix.
4. Return GATEWAY™ BP Clonase™ Enzyme Mix to - 80° C freezer.
5. Incubate tubes at 25° for at least 60 minutes. If both the attB and attP DNAs are supercoiled, incubation for 2-24 hours at 25°C is recommended.
6. Add 2 μ l Clonase Stop Solution. Incubate for 10 min at 37°C.
7. Transform 2 μ l into 100 μ l competent E. coli, as above. Select on LB plates containing 50 μ g/ml kanamycin.

Example 11: Cloning PCR Products into Entry Vectors using Standard Cloning Methods

Preparation of Entry Vectors for Cloning of PCR Products

All of the Entry Vectors of the invention contain the death gene ccdB as a stuffer between the "left" and "right" restriction sites. The advantage of this arrangement is that there is virtually no background from vector that has not been cut with both restriction enzymes, because the presence of the ccdB gene will kill

all standard E. coli strains. Thus it is necessary to cut each Entry Vector twice, to remove the ccdB fragment.

We strongly recommend that, after digestion of the Entry Vector with the second restriction enzyme, you treat the reaction with phosphatase (calf intestine alkaline phosphatase, CIAP or thermosensitive alkaline phosphatase, TSAP). The phosphatase can be added directly to the reaction mixture, incubated for an additional time, and inactivated. This step dephosphorylates both the vector and ccdB fragments, so that during subsequent ligation there is less competition between the ccdB fragment and the DNA of interest for the termini of the Entry Vector.

Blunt Cloning of PCR products

Generally PCR products do not have 5' phosphates (because the primers are usually 5' OH), and they are not necessarily blunt. (On this latter point, see Brownstein, et al., *BioTechniques* 20: 1006, 1996 for a discussion of how the sequence of the primers affects the addition of single 3' bases.) The following protocol repairs these two defects.

In a 0.5 ml tube, ethanol precipitate about 40 ng of PCR product (as judged from an agarose gel).

1. Dissolve the precipitated DNA in 10 μ l comprising 1 μ l 10 mM rATP, 1 μ l mixed 2 mM dNTPs (i.e., 2 mM each dATP, dCTP, dTTP, and dGTP), 2 μ l 5x T4 polynucleotide kinase buffer (350 mM Tris HCl (pH7.6), 50 mM $MgCl_2$, 500mM KCl, 5 mM 2-mercaptoethanol) 10 units T4 polynucleotide kinase, 1 μ l T4 DNA polymerase, and water to 10 μ l.
2. Incubate the tube at 37° for 10 minutes, then at 65° for 15 minutes, cool, centrifuge briefly to bring any condensate to the tip of the tube.
3. Add 5 μ l of the PEG/ $MgCl_2$ solution, mix and centrifuge at room temperature for 10 minutes. Discard supernatant.
4. Dissolve the invisible precipitate in 10 μ l containing 2 μ l 5x T4 DNA ligase buffer (Life Technologies, Inc.), 0.5 units T4 DNA ligase, and about 50 ng of blunt, phosphatase-treated Entry Vector.

5. Incubate at 25° for 1 hour, then 65° for 10 minutes. Add 90 µl TE, transform 10 µl into 50 - 100 µl competent *E. coli* cells.
6. Plate on kanamycin.

5 **Note:** In the above protocol, steps b-c simultaneously polish the ends of the PCR product (through the exonuclease and polymerase activities of T4 DNA polymerase) and phosphorylate the 5' ends (using T4 polynucleotide kinase). It is necessary to inactivate the kinase, so that the blunt, dephosphorylated vector in step e cannot self ligate. Step d (the PEG precipitation) removes all small
10 molecules (primers, nucleotides), and has also been found to improve the yield of cloned PCR product by 50 fold.

Cloning PCR Products after Digestion with Restriction Enzymes

Efficient cloning of PCR products that have been digested with restriction
15 enzymes includes three steps: inactivation of *Taq* DNA polymerase, efficient restriction enzyme cutting, and removal of small DNA fragments.

Inactivation of Taq DNA Polymerase: Carryover of *Taq* DNA polymerase and dNTPs into a RE digestion significantly reduces the success in cloning a PCR product (D. Fox et al., *FOCUS* 20(1):15, 1998), because *Taq* DNA polymerase
20 can fill in sticky ends and add bases to blunt ends. Either TAQQUENCH™ (obtainable from Life Technologies, Inc.; Rockville, Maryland) or extraction with phenol can be used to inactivate the *Taq*.

Efficient Restriction Enzyme Cutting: Extra bases on the 5' end of each PCR primer help the RE cut near ends of PCR products. With the availability of
25 cheap primers, adding 6 to 9 bases on the 5' sides of the restriction sites is a good investment to ensure that most of the ends are digested. Incubation of the DNA with a 5-fold excess of restriction enzyme for an hour or more helps ensure success.

Removal of Small Molecules before Ligation: Primers, nucleotides,
30 primer dimers, and small fragments produced by the restriction enzyme digestion,

can all inhibit or compete with the desired ligation of the PCR product to the cloning vector. This protocol uses PEG precipitation to remove small molecules.

Protocol for cutting the ends of PCR products with restriction enzyme(s):

1. Inactivation of Taq DNA polymerase in the PCR product:

Option A: Extraction with Phenol

A1. Dilute the PCR reaction to 200 μ l with TE. Add an equal volume of phenol:chloroform:isoamyl alcohol, vortex vigorously for 20 seconds, and centrifuge for 1 minute at room temperature. Discard the lower phase.

A2. Extract the phenol from the DNA and concentrate as follows. Add an equal volume of 2-butanol (colored red with "Oil Red O" from Aldrich, if desired), vortex briefly, centrifuge briefly at room temperature. Discard the upper butanol phase. Repeat the extraction with 2-butanol. This time the volume of the lower aqueous phase should decrease significantly. Discard the upper 2-butanol phase.

A3. Ethanol precipitate the DNA from the aqueous phase of the above extractions. Dissolve in a 200 μ l of a suitable restriction enzyme (RE) buffer.

Option B: Inactivation with TaqQuench

B1. Ethanol precipitate an appropriate amount of PCR product (100 ng to 1 μ g), dissolve in 200 μ l of a suitable RE buffer.

B2. Add 2 μ l TaqQuench.

2. Add 10 to 50 units of restriction enzyme and incubate for at least 1 hour. Ethanol precipitate if necessary to change buffers for digestion at the other end of the PCR product.

3. Add ½ volume of the PEG/MgCl₂ mix to the RE digestion. Mix well and immediately centrifuge at room temperature for 10 minutes. Discard the supernatant (pellet is usually invisible), centrifuge again for a few seconds, discard any remaining supernatant.

4. Dissolve the DNA in a suitable volume of TE (depending on the amount of PCR product in the original amplification reaction) and apply an aliquot to an agarose gel to confirm recovery. Apply to the same gel 20-100 ng of the appropriate Entry Vector that will be used for the cloning.

Example 12: Determining The Expected Size of the GATEWAY™ Cloning Reaction Products

If you have access to a software program that will electronically cut and splice sequences, you can create electronic clones to aid you in predicting the sizes and restriction patterns of GATEWAY™ Cloning System recombination products.

The cleavage and ligation steps performed by the enzyme Int in the GATEWAY™ Cloning System recombination reactions mimic a restriction enzyme cleavage that creates a 7-bp 5'-end overhang followed by a ligation step that reseals the ends of the daughter molecules. The recombination proteins present in the Clonase cocktails (see Example 19 below) recognize the 15 bp core sequence present within all four types of att sites (in addition to other flanking sequences characteristic of each of the different types of att sites).

By treating these sites in your software program as if they were restriction sites, you can cut and splice your Entry Clones with various Destination Vectors and obtain accurate maps and sequences of the expected results from your GATEWAY™ Cloning System reactions.

Example 13: Protein Expression

Brief Review of Protein Expression

Transcription: The most commonly used promoters in *E. coli* Expression Vectors are variants of the lac promoter, and these can be turned on by adding

IPTG to the growth medium. It is usually good to keep promoters off until expression is desired, so that the host cells are not made sick by the overabundance of some heterologous protein. This is reasonably easy in the case of the lac promoters used in *E. coli*. One needs to supply the *lac I* gene (or its more productive relative, the *lac I^s* gene) to make *lac* repressor protein, which binds near the promoter and keeps transcription levels low. Some Destination Vectors for *E. coli* expression carry their own *lacI^s* gene for this purpose. (However, lac promoters are always a little "on," even in the absence of IPTG.)

Controlling transcription in eukaryotic cells is not nearly so straightforward or efficient. The tetracycline system of Bujard and colleagues is the most successful approach, and one of the Destination Vectors (pDEST11; Figure 31) has been constructed to supply this function.

Translation: Ribosomes convert the information present in mRNA into protein. Ribosomes scan RNA molecules looking for methionine (AUG) codons, which begin nearly all nascent proteins. Ribosomes must, however, be able to distinguish between AUG codons that code for methionine in the middle of proteins from those at the start. Most often ribosomes choose AUGs that are 1) first in the RNA (toward the 5' end), and 2) have the proper sequence context. In *E. coli* the favored context (first recognized by Shine and Dalgarno, *Eur. J. Biochem.* 57: 221 (1975)) is a run of purines (As and Gs) from five to 12 bases upstream of the initiating AUG, especially AGGAGG or some variant.

In eukaryotes, a survey of translated mRNAs by Kozak (*J. Biol. Chem.* 266: 19867 (1991)) has revealed a preferred sequence context, gcc Acc ATGG, around the initiating methionine, with the A at -3 being most important, and a purine at +4 (where the A of the ATG is +1), preferably a G, being next most influential. Having an A at -3 is enough to make most ribosomes choose the first AUG of an mRNA, in plants, insects, yeast, and mammals. (For a review of initiation of protein synthesis in eukaryotic cells, see: Pain, V.M. *Eur. J. Biochem.* 236:747-771, 1996.)

Consequences of Translation Signals for GATEWAY™ Cloning System: First, translation signals (Shine-Dalgarno in *E. coli*, Kozak in eukaryotes) have to be close to the initiating ATG. The attB site is 25 base pairs long. Thus if

translation signals are desired near the natural ATG of the nucleic acid molecule of interest, they must be present in the Entry Clone of that nucleic acid molecule of interest. Also, when a nucleic acid molecule of interest is moved from an Entry Clone to a Destination vector, any translation signals will move along. The result is that the presence or absence of Shine-Dalgarno and/or Kozak sequences in the Entry Clone must be considered, with the eventual Destination Vectors to be used in mind.

Second, although ribosomes choose the 5' ATG most often, internal ATGs are also used to begin protein synthesis. The better the translation context around this internal ATG, the more internal translation initiation will be seen. This is important in the GATEWAY™ Cloning System, because you can make an Entry Clone of your nucleic acid molecule of interest, and arrange to have Shine-Dalgarno and/or Kozak sequences near the ATG. When this cassette is recombined into a Destination Vector that transcribes your nucleic acid molecule of interest, you get native protein. If you want, you can make a fusion protein in a different Destination Vector, since the Shine-Dalgarno and/or Kozak sequences do not contain any stop signals in the same reading frame. However, the presence of these internal translation signals may result in a significant amount of native protein being made, contaminating, and lowering the yield of, your fusion protein. This is especially likely with short fusion tags, like His6.

A good compromise can be recommended. If an Entry Vector like pENTR7 (Figure 16) or pENTR8 (Figure 17) is chosen, the Kozak bases are present for native eukaryotic expression. The context for *E. coli* translation is poor, so the yield of an amino-terminal fusion should be good, and the fusion protein can be digested with the TEV protease to make near-native protein following purification.

Recommended Conditions for Synthesis of Proteins in E. coli: When making proteins in *E. coli* it is advisable, at least initially, to incubate your cultures at 30°C, instead of at 37°C. Our experience indicates that proteins are less likely to form aggregates at 30°C. In addition, the yields of proteins from cells grown at 30°C frequently are improved.

The yields of proteins that are difficult to express may also be improved by inducing the cultures in mid-log phase of growth, using cultures begun in the morning from overnight growths, as opposed to harvesting directly from an overnight culture. In the latter case, the cells are preferably in late log or stationary growth, which can favor the formation of insoluble aggregates.

Example 14: Constructing Destination Vectors from Existing Vectors

Destination Vectors function because they have two recombination sites, attR1 and attR2, flanking a chloramphenicol resistance (CmR) gene and a death gene, ccdB. The GATEWAY™ Cloning System recombination reactions exchange the entire Cassette (except for a few bases comprising part of the attB sites) for the DNA segment of interest from the Entry Vector. Because attR1, CmR, ccdB gene, and attR2 are contiguous, they can be moved on a single DNA segment. If this Cassette is cloned into a plasmid, the plasmid becomes a Destination Vector. Figure 63 shows a schematic of the GATEWAY™ Cloning System Cassette; attR cassettes in all three reading frames contained in vectors pEYC15101, pEYC15102 and pEYC15103 are shown in Figures 64A, 64B, and 64C, respectively.

The protocol for constructing a Destination Vector is presented below. Keep in mind the following points:

- Destination Vectors must be constructed and propagated in one of the DB strains of *E. coli* (e.g., DB3.1, and particularly *E. coli* LIBRARY EFFICIENCY® DB3.1™ Competent Cells) available from Life Technologies, Inc. (and described in detail in U.S. Provisional Application No. 60/122,392, filed on March 2, 1999, which is incorporated herein by reference), because the ccdB death gene will kill any *E. coli* strain that has not been mutated such that it will survive the presence of the ccdB gene.
- If your Destination Vector will be used to make a fusion protein, a GATEWAY™ Cloning System cassette with the correct reading frame must be used. The nucleotide sequences of the ends of the cassettes are shown in Figure 78. The reading frame of the fusion protein domain must

be in frame with the core region of the attR1 site (for an amino terminal fusion) so that the six As are translated into two lysine codons. For a C-terminal fusion protein, translation through the core region of the attR2 site should be in frame with -TAC-AAA-, to yield -Tyr-Lys-.

- Note that each reading frame Cassette has a different unique restriction site between the chloramphenicol resistance and ccdB genes (*Mlu*I for reading frame A, *Bgl*II for reading frame B, and *Xba*I for reading frame C; see Figure 63).
- Most standard vectors can be converted to Destination Vectors, by inserting the Entry Cassette into the MCS of that vector.

Protocol for Making a Destination Vector

1. If the vector will make an amino fusion protein, it is necessary to keep the "aaa" triplets in attR1 in phase with the triplets of the fusion protein. Determine which Entry cassette to use as follows:

a.) Write out the nucleotide sequence of the existing vector near the restriction site into which the Entry cassette will be cloned. These must be written in triplets corresponding to the amino acid sequence of the fusion domain.

b.) Draw a vertical line through the sequence that corresponds to the restriction site end, after it has been cut and made blunt, i.e., after filling in a protruding 5' end or polishing a protruding 3' end.

c.) Choose the appropriate reading frame cassette:

- If the coding sequence of the blunt end ends after a complete codon triplet, use the reading frame A cassette. See Figures 78, 79 and 80.

- If the coding sequence of the blunt end ends in a single base, use the reading frame B cassette. See Figures 78, 79 and 81.

- If the coding sequence of the blunt end ends in two bases, use the reading frame C cassette. See Figures 78, 79, 82A-B, and 83A-C.

2. Cut one to five micrograms of the existing plasmid at the position where you wish your nucleic acid molecule of interest (flanked by att sites) to be after the recombination reactions. Note: it is better to remove as many of the MCS restriction sites as possible at this step. This makes it more likely that restriction enzyme sites within the GATEWAY™ Cloning System Cassette will be unique in the new plasmid, which is important for linearizing the Destination Vector (Example 14, below).

3. Remove the 5' phosphates with alkaline phosphatase. While this is not mandatory, it increases the probability of success.

4. Make the end(s) blunt with fill-in or polishing reactions. For example, to 1 µg of restriction enzyme-cut, ethanol-precipitated vector DNA, add:

- 20 µl 5x T4 DNA Polymerase Buffer (165 mM Tris-acetate (pH 7.9), 330 mM Na acetate, 50 mM Mg acetate, 500 µg/ml BSA, 2.5 mM DTT)
- 5 µl 10mM dNTP mix
- 1 Unit of T4 DNA Polymerase
- Water to a final volume of 100 µl
- Incubate for 15 min at 37°C.

5. Remove dNTPs and small DNA fragments: Ethanol precipitate (add three volumes of room temperature ethanol containing 0.1 M sodium acetate, mix well, immediately centrifuge at room temperature 5 - 10 minutes), dissolve wet precipitate in 200 µl TE, add 100 µl 30% PEG 8000, 30 mM MgCl₂, mix well,

immediately centrifuge for 10 minutes at room temperature, discard supernatant, centrifuge again a few seconds, discard any residual liquid.

6. Dissolve the DNA to a final concentration of 10 - 50 ng per microliter. Apply
20 - 100 ng to a gel next to supercoiled plasmid and linear size standards to confirm cutting and recovery. The cutting does not have to be 100% complete, since you will be selecting for the chloramphenicol marker on the Entry cassette.

7. In a 10 μ l ligation reaction combine 10 - 50 ng vector, 10 - 20 ng of Entry Cassette (Figure 79), and 0.5 units T4 DNA ligase in ligase buffer. After one hour (or overnight, whichever is most convenient), transform 1 μ l into one of the DB strains of competent *E. coli* cells with a *gyrA462* mutation (See U.S. Provisional Application No. 60/122,392, filed on March 2, 1999, which is incorporated herein by reference), preferably DB3.1, and most preferably *E. coli* LIBRARY EFFICIENCY® DB3.1™ Competent Cells. The *ccdB* gene on the Entry Cassette will kill other strains of *E. coli* that have not been mutated so as to survive the presence of the *ccdB* gene.

8. After expression in SOC medium, plate 10 μ l and 100 μ l on chloramphenicol-containing (30 μ g / ml) plates, incubate at 37° C.

9. Pick colonies, make miniprep DNA. Treat the miniprep with RNase A and store in TE. Cut with the appropriate restriction enzyme to determine the orientation of the Cassette. Choose clones with the attR1 site next to the amino end of the protein expression function of the plasmid.

Notes on Using Destination Vectors

- We have found that about ten-fold more colonies result from a GATEWAY™ Cloning System reaction if the Destination Vector is linear or relaxed. If the competent cells you use are highly competent ($>10^8$ per microgram), linearizing the Destination Vector is less essential.

- The site or sites used for the linearization must be within the Entry Cassette. Sites that cut once or twice within each cassette are shown in Figures 80-82.
- Minipreps of Destination Vectors will work fine, so long as they have been treated with RNase. Since most DB strains are *endA*- (See U.S. Provisional Application No. 60/122,392, filed on March 2, 1999, which is incorporated herein by reference), minipreps can be digested with restriction enzymes without a prior phenol extraction.
- Reading the OD₂₆₀ of miniprep DNA is inaccurate unless the RNA and ribonucleotides have been removed, for example, by a PEG precipitation.

Example 15: Some Options in Choosing Appropriate Entry Vectors and Destination Vectors: An Example

In some applications, it may be desirable to express a nucleic acid molecule of interest in two forms: as an amino-terminal fusion in *E. coli*, and as a native protein in eukaryotic cells. This may be accomplished in any of several ways:

Option 1: Your choices depend on your nucleic acid molecule of interest and the fragment that contains it, as well as the available Entry Vectors. For eukaryotic translation, you need consensus bases according to Kozak (*J. Biol. Chem.* 266:19867, 1991) near the initiating methionine (ATG) codon. All of the Entry Vectors offer this motif upstream of the *Xmn*I site (blunt cutter). One option is to amplify your nucleic acid molecule of interest, with its ATG, by PCR, making the amino end blunt and the carboxy end containing the natural stop codon followed by one of the "right side" restriction sites (*Eco*RI, *Not*I, *Xho*I, *Eco*RV, or *Xba*I of the pENTR vectors).

If you know your nucleic acid molecule of interest does not have, for example, an *Xho*I site, you can make a PCR product that has this structure:

Xho I

5' ATG nnn nnn --- nnn TAA ctc gag nnn nnn 3'
 3' tac nnn nnn --- nnn att gag ctc nnn nnn 5'

After cutting with *Xho*I, the fragment is ready to clone:

```
5' ATG nnn nnn --- nnn TAA c      3'
3' tac nnn nnn --- nnn att gag ct  5'
```

(If you follow this example, don't forget to put a phosphate on the amino oligo.)

Option 2: This PCR product could be cloned into two Entry Vectors to give the desired products, between the *Xmn*I and *Xho*I sites: pENTR1A (Figures 10A, 10B) or pENTR7 (Figures 16A, 16B). If you clone into pENTR1A, amino fusions will have the minimal number of amino acids between the fusion domain and your nucleic acid molecule of interest, but the fusion cannot be removed with TEV protease. The converse is true of clones in pENTR7, i.e., an amino fusion can be cleaved with TEV protease, at the cost of more amino acids between the fusion and your nucleic acid molecule of interest.

In this example, let us choose to clone our hypothetical nucleic acid molecule of interest into pENTR7, between the *Xmn*I and *Xho*I sites. Once this is accomplished, several optional protocols using the Entry Clone pENTR7 may be followed:

Option 3: Since the nucleic acid molecule of interest has been amplified with PCR, it may be desirable to sequence it. To do this, transfer the nucleic acid molecule of interest from the Entry Vector into a vector that has priming sites for the standard sequencing primers. Such a vector is pDEST6 (Figures 26A, 26B). This Destination Vector places the nucleic acid molecule of interest in the opposite orientation to the lac promoter (which is leaky -- see Example 3 above). If the gene product is toxic to *E. coli*, this Destination Vector will minimize its toxicity.

Option 4: While the sequencing is going on, you might wish to check the expression of the nucleic acid molecule of interest in, for example, CHO cells, by recombining the nucleic acid molecule of interest into a CMV promoter vector (pDEST7, Figure 27; or pDEST12, Figure 32), or into a baculovirus vector (pDEST8, Figure 28; or pDEST10, Figure 30) for expression in insect cells. Both

of these vectors will transcribe the coding sequence of your nucleic acid molecule of interest, and translate it from the ATG of the PCR product using the Kozak bases upstream of the *Xmn*I site.

Option 5: If you wish to purify protein, for example to make antibodies, you can clone the nucleic acid molecule of interest into a His6 fusion vector, pDEST2 (Figure 22). Since the nucleic acid molecule of interest is cloned downstream of the TEV protease cleavage domain of pENTR7 (Figure 16), the amino acid sequence of the protein produced will be:

[----- attB1 -----] TEV protease
 NH2- MSYYHHHHHHGITSLYKKAGFENLYFQ! GTM----COOH

The attB site and the restriction sites used to make the Destination and Entry Vectors are translated into the underlined 11 amino acids (GITSLYKKAGF). Cleavage with TEV protease (arrow) leaves two amino acids, GT, on the amino end of the gene product.

See Figure 55 for an example of a nucleic acid molecule of interest, the chloramphenicol acetyl transferase (CAT) gene, cloned into pENTR7 (Figure 16) as a blunt (amino)-*Xho*I (carboxy) fragment, then cloned by recombination into the His6 fusion vector pDEST2 (Figure 22).

Option 6: If the His6 fusion protein is insoluble, you may go on and try a GST fusion. The appropriate Destination vector is pDEST3 (Figure 23).

Option 7: If you need to make RNA probes and prefer SP6 RNA polymerase, you can make the top strand RNA with your nucleic acid molecule of interest, cloned into pSPORT+ (pDEST5 (Figures 25A, 25B)), and the bottom strand RNA with the nucleic acid molecule of interest cloned into pSPORT(-) (pDEST6 (Figures 26A, 26B)). Opposing promoters for T7 RNA polymerase and SP6 RNA polymerase are also present in these clones.

Option 8: It is often worthwhile to clone your nucleic acid molecule of interest into a variety of Destination Vectors in the same experiment. For example, if the number of colonies varies widely when the various recombination reactions are transformed into *E. coli*, this may be an indication that the nucleic acid molecule of interest is toxic in some contexts. (This problem is more clearly evident when a positive control gene is used for each Destination Vector.) Specifically, if many more colonies are obtained when the nucleic acid molecule of interest is recombined into pDEST6 than in pDEST5, there is a good chance that leakiness of the lac promoter is causing some expression of the nucleic acid molecule of interest in pSPORT "+" (which is not harmful in pDEST6 because the nucleic acid molecule of interest is in the opposite orientation).

Example 16: Demonstration of a One-tube Transfer of a PCR Product (or Expression Clone) to Expression Clone via a Recombinational Cloning Reaction

In the BxP recombination (Entry or Gateway) reaction described herein, a DNA segment flanked by attB1 and attB2 sites in a plasmid conferring ampicillin resistance was transferred by recombination into an attP plasmid conferring kanamycin resistance, which resulted in a product molecule wherein the DNA segment was flanked by attL sites (attL1 and attL2). This product plasmid comprises an "attL Entry Clone" molecule, because it can react with a "attR Destination Vector" molecule via the LxR (Destination) reaction, resulting in the transfer of the DNA segment to a new (ampicillin resistant) vector. In the previously described examples, it was necessary to transform the BxP reaction products into *E. coli*, select kanamycin resistant colonies, grow those colonies in liquid culture, and prepare miniprep DNA, before reacting this DNA with a Destination Vector in an LxR reaction.

The goal of the following experiment was to eliminate the transformation and miniprep DNA steps, by adding the BxP Reaction products directly to an LxR Reaction. This is especially appropriate when the DNA segment flanked by attB sites is a PCR product instead of a plasmid, because the PCR product cannot give

ampicillin-resistant colonies upon transformation, whereas attB plasmids (in general) carry an ampicillin resistance gene. Thus use of a PCR product flanked by attB sites in a BxP Reaction allows one to select for the ampicillin resistance encoded by the desired attB product of a subsequent LxR Reaction.

5 Two reactions were prepared: Reaction A, negative control, no attB PCR product, (8 µl) contained 50 ng pEZC7102 (attP Donor plasmid, confers kanamycin resistance) and 2 µl BxP Clonase (22 ng / µl Int protein and 8 ng/µl IHF protein) in BxP buffer (25 mM Tris HCl, pH 7.8, 70 mM KCl, 5 mM spermidine, 0.5 mM EDTA, 250 µg / ml BSA). Reaction B (24 µl) contained
10 150 ng pEZC7102, 6 µl BxP Clonase, and 120 ng of the attB -tet-PCR product in the same buffer as reaction A. The attB - tet - PCR product comprised the tetracycline resistance gene of plasmid pBR322, amplified with two primers containing either attB1 or attB2 sites, and having 4 Gs at their 5' ends, as described earlier.

15 The two reactions were incubated at 25°C for 30 minutes. Then aliquots of these reactions were added to new components that comprised LxR Reactions or appropriate controls for the LxR Reaction. Five new reactions were thus produced:

20 **Reaction 1:** 5 µl of reaction A was added to a 5 µl LxR Reaction containing 25 ng NcoI-cut pEZC8402 (the attR Destination Vector plasmid) in LxR buffer (37.5 mM Tris HCl, pH 7.7, 16.5 mM NaCl, 35 mM KCl, 5 mM spermidine, 375 µg / ml BSA), and 1 µl of GATEWAY™ LR Clonase™ Enzyme Mix (total volume of 10 µl).

25 **Reaction 2:** Same as reaction 1, except 5 µl of reaction B (positive) were added instead of reaction A (negative).

30 **Reaction 3:** Same as reaction 2, except that the amounts of Nco-cut pEZC8402 and GATEWAY™ LR Clonase™ Enzyme Mix were doubled, to 50 ng and 2 µl, respectively.

Reaction 4: Same as reaction 2, except that 25 ng of pEZ11104 (a positive control attL Entry Clone plasmid) were added in addition to the aliquot of reaction B.

Reaction 5: Positive control LxR Reaction, containing 25 ng *Nco*I-cut pEZC8402, 25 ng pEZ11104, 37.5 mM Tris HCl pH 7.7, 16.5 mM NaCl, 35 mM KCl, 5 mM spermidine, 375 µg / ml BSA and 1 µl GATEWAY™ LR Clonase™ Enzyme Mix in a total volume of 5 µl.

All five reactions were incubated at 25°C for 30 minutes. Then, 1 µl aliquots of each of the above five reactions, plus 1 µl from the remaining volume of Reaction B, the standard BxP Reaction, were used to transform 50 µl competent DH5α *E. coli*. DNA and cells were incubated on ice for 15 min., heat shocked at 42°C for 45 sec., and 450 µl SOC were added. Each tube was incubated with shaking at 37°C for 60 min. Aliquots of 100 µl and 400 µl of each transformation were plated on LB plates containing either 50 µg/ml kanamycin or 100 µg/ml ampicillin (see Table 2). A transformation with 10 pg of pUC19 DNA (plated on LB-amp₁₀₀) served as a control on the transformation efficiency of the DH5α cells. Following incubation overnight at 37°C, the number of colonies on each plate was determined.

Results of these reactions are shown in Table 2.

Table 2*

Reaction No.:	1	2	3	4	5	6
	Number of Colonies					
Vol. plated:	Neg. Control BxP Reaction	1X pEZC8402 and LR Clonase™	2X pEZC8402 and LR Clonase™	LxR Reaction with Pos. Control DNA	LxR Reaction alone	BxP Reaction alone
100 µl	2	1	8	9	~1000	~1000
400 µl	5	10	35	62	>2000	>2000
Selection:	Kan	Amp	Amp	Amp	Amp	Kan

*(Transformation with pUC 19 DNA yielded 1.4×10^9 CFU/µg DNA.)

34 of the 43 colonies obtained from Reaction 3 were picked into 2 ml Terrific Broth with 100 µg/ml ampicillin and these cultures were grown overnight, with shaking, at 37°C. 27 of the 34 cultures gave at least moderate growth, and of these 24 were used to prepare miniprep DNA, using the standard protocol. These 24 DNAs were initially analyzed as supercoiled (SC) DNA on a 1% agarose gel to identify those with inserts and to estimate the sizes of the inserts. Fifteen of the 24 samples displayed SC DNA of the size predicted (5553 bp) if **tetx7102** had correctly recombined with **pEJC8402** to yield **tetx8402**. One of these samples contained two plasmids, one of ~5500 bp and a one of ~3500 bp. The majority of the remaining clones were approximately 4100 bp in size.

All 15 of the clones displaying SC DNA of predicted size (~5500 bp) were analyzed by two different double digests with restriction endonucleases to confirm the structure of the expected product: **tetx8402**. (See plasmid maps, Figures 57-59) In one set of digests, the DNAs were treated with Not I and Eco RI, which should cut the predicted product just outside both attB sites, releasing the tet' insert on a fragment of 1475 bp. In the second set of digests, the DNAs were digested with *NotI* and with *NruI*. *NruI* cleaves asymmetrically within the subcloned tet' insert, and together with *NotI* will release a fragment of 1019 bp.

Of the 15 clones analyzed by double restriction digestion, 14 revealed the predicted sizes of fragments for the expected product.

Interpretation:

The DNA components of Reaction B, pEJC7102 and attB-tet-PCR, are shown in Figure 56. The desired product of BxP Reaction B is **tetx7102**, depicted in Figure 57. The LxR Reaction recombines the product of the BxP Reaction, **tetx7102** (Figure 57), with the Destination Vector, pEJC8402, shown in Figure 58. The LxR Reaction with **tetx7102** plus pEJC8402 is predicted to yield the desired product **tetx8402**, shown in Figure 59.

Reaction 2, which combined the BxP Reaction and LxR Reaction, gave few colonies beyond those of the negative control Reaction. In contrast, Reaction 3, with twice the amount of pEJC8402 (Figure 58) and LxR Clonase, yielded a

larger number of colonies. These colonies were analyzed further, by restriction digestion, to confirm the presence of expected product. Reaction 4 included a known amount of attL Entry Clone plasmid in the combined BxP-plus-LxR reaction. But reaction 4 yielded only about 1% of the colonies obtained when the same DNA was used in a LxR reaction alone, Reaction 6. This result suggests that the LxR reaction may be inhibited by components of the BxP reaction.

Restriction endonuclease analysis of the products of Reaction 3 revealed that a sizeable proportion of the colonies (14 of the 34 analyzed) contained the desired tet^r subclone, tetx8402 (Figure 59).

The above results establish the feasibility of performing first a BxP recombination reaction followed by a LxR recombination reaction -- in the same tube -- simply by adding the appropriate buffer mix, recombination proteins, and DNAs to a completed BxP reaction. This method should prove useful as a faster method to convert attB-containing PCR products into different Expression Clones, eliminating the need to isolate first the intermediate attL-PCR insert subclones, before recombining these with Destination Vectors. This may prove especially valuable for automated applications of these reactions.

This same one-tube approach allows for the rapid transfer of nucleic acid molecules contained in attB plasmid clones into new functional vectors as well. As in the above examples, attL subclones generated in a BxP Reaction can be recombined directly with various Destination Vectors in a LxR reaction. The only additional requirement for using attB plasmids, instead of attB-containing PCR products, is that the Destination Vector(s) employed must contain a different selection marker from the one present on the attB plasmid itself and the attP vector.

Two alternative protocols for a one-tube reaction have also proven useful and somewhat more optimal than the conditions described above.

Alternative 1:

Reaction buffer contained 50 mM Tris-HCl (pH 7.5), 50 mM NaCl, 0.25 mM EDTA, 2.5 mM spermidine, and 200 µg/ml BSA. After a 16 (or 3) hour incubation of the PCR product (100 ng) + attP Donor plasmid (100 ng) +

GATEWAY™ BP Clonase™ Enzyme Mix + Destination Vector (100 ng), 2 µl of GATEWAY™ LR Clonase™ Enzyme Mix (per 10 µl reaction mix) was added and the mixture was incubated an additional 6 (or 2) hours at 25°C. Stop solution was then added as above and the mixture was incubated at 37°C as above and transformed by electroporation with 1 µl directly into electrocompetent host cells. Results of this series of experiments demonstrated that longer incubation times (16 hours vs. 3 hours for the BP Reaction, 6 hours vs. 2 hours for the LR Reaction) resulted in about twice as many colonies being obtained as for the shorter incubation times. With two independent genes, 10/10 colonies having the correct cloning patterns were obtained.

Alternative 2:

A standard BP Reaction under the reaction conditions described above for Alternative 1 was performed for 2 hours at 25°C. Following the BP Reaction, the following components were added to the reaction mixture in a total volume of 7 µl:

20 mM Tris-HCl, pH 7.5

100 mM NaCl

5 µg/ml Xis-His6

15% glycerol

~1000 ng of Destination Vector

The reaction mixture was then incubated for 2 hours at 25°C, and 2.5 µl of stop solution (containing 2 µg/ml proteinase K) was added and the mixture was incubated at 37°C for an additional 10 minutes. Chemically competent host cells were then transformed with 2 µl of the reaction mixture, or electrocompetent host cells (e.g., EMax DH10B cells; Life Technologies, Inc.) were electroporated with 2 µl of the reaction mixture per 25-40 µl of cells. Following transformation, mixtures were diluted with SOC, incubated at 37°C, and plated as described above on media selecting for the selection markers on the Destination Vector and the Entry clone (B x P reaction product). Analogous results to those described for Alternative 1 were obtained with these reaction conditions -- a higher level of colonies containing correctly recombined reaction products were observed.

Example 17: Demonstration of a One-tube Transfer of a PCR Product (or Expression Clone) to Expression Clone via a Recombinational Cloning Reaction

5 Single-tube transfer of PCR product DNA or Expression Clones into Expression Clones by recombinational cloning has also been accomplished using a procedure modified from that described in Example 16. This procedure is as follows:

10 •Perform a standard BP (Gateway) Reaction (see Examples 9 and 10) in 20 µl volume at 25°C for 1 hour.

15 •After the incubation is over, take a 10 µl aliquot from the 20 µl total volume and add 1 µl of Proteinase K (2 mg/ml) and incubate at 37°C for 10 minutes. This first aliquot can be used for transformation and gel assay of BP reaction analysis. Plate BP reaction transformation on LB plates with **Kanamycin** (50 µg/ml).

20 •Add the following reagents to the remaining 10 µl aliquot of the BP reaction:

1 µl of 0.75 M NaCl

2 µl of destination vector (150 ng/µl)

4 µl of LR Clonase™ (after thawing and brief mixing)

25 •Mix all reagents well and incubate at 25°C for 3 hours. Stop the reaction at the end of incubation with 1.7 µl of Proteinase K (2 mg/ml) and incubate at 37°C for 10 minutes.

30 •Transform 2 µl of the completed reaction into 100 µl of competent cells. Plate 100 µl and 400 µl on LB plates with **Ampicillin** (100 µg/ml).

Notes:

•If your competent cells are less than 10⁸ CFU/µg, and you are concerned about getting enough colonies, you can improve the yield several fold by incubating the

BP reaction for 6-20 hours. Electroporation also can yield better colony output than chemical transformation.

•PCR products greater than about 5-6 kb show significantly lower cloning efficiency in the BP reaction. In this case, we recommend using longer incubation times for both BP and LR steps.

•If you want to move your insert gene into several destination vectors simultaneously, then scale up the initial BP reaction volume so that you have a 10 µl aliquot for adding each destination vector.

Example 18: Optimization of GATEWAY™ Clonase™ Enzyme Compositions

The enzyme compositions containing Int and IHF (for BP Reactions) were optimized using a standard functional recombinational cloning reaction (a BP reaction) between attB-containing plasmids and attP-containing plasmids, according to the following protocol:

Materials and Methods:

Substrates:

AttP - supercoiled pDONR201

AttB - linear ~ 1Kb [³H]PCR product amplified from pEZC7501

Proteins:

IntH6 -- His₆-carboxy- tagged λ Integrase

IHF -- Integration Host Factor

Clonase:

50 ng/µl IntH6 and 20 ng/µl IHF, admixed in 25 mM Tris- HCl (pH 7.5), 22 mM NaCl, 5 mM EDTA, 1 mg/ml BSA, 5 mM Spermidine, and 50% glycerol.

Reaction Mixture (total volume of 40 μ l):

1000 ng AttP plasmid

600 ng AttB [3 H] PCR product8 μ l Clonase (400 ng IntH6, 160 ng IHF) in 25 mM Tris-HCl (pH 7.5),

22 mM NaCl, 5 mM EDTA, 1 mg/ml BSA, 5 mM Spermidine, 5 mM

DTT.

Reaction mixture was incubated for 1 hour at 25°C, 4 μ l of 2 μ g/ μ l proteinase K was added and mixture was incubated for an additional 20 minutes at 37°C. Mixture was then extracted with an equal volume of Phenol/Chloroform/Isoamyl alcohol. The aqueous layer was then collected, and 0.1 volumes of 3 M sodium acetate and 2 volumes of cold 100% ethanol were added. Tubes were then spun in a microcentrifuge at maximum RPM for 10 minutes at room temperature. Ethanol was decanted, and pellets were rinsed with 70% ethanol and re-centrifuged as above. Ethanol was decanted, and pellets were allowed to air dry for 5-10 minutes and then dissolved in 20 μ l of 33 mM Tris-Acetate (pH 7.8), 66 mM potassium acetate, 10 mM magnesium acetate, 1 mM DTT, and 1mM ATP. 2 units of exonuclease V (e.g., Plasmid Safe; EpiCentre, Inc., Madison, WI) was then added, and the mixture was incubated at 37°C for 30 minutes.

Samples were then TCA-washed by spotting 30 μ l of reaction mixture onto a Whatman GF/C filter, washing filters once with 10% TCA + 1% NaPPi for 10 minutes, three times with 5% TCA for 5 minutes each, and twice with ethanol for 5 minutes each. Filters were then dried under a heat lamp, placed into a scintillation vial, and counted on a β liquid scintillation counter (LSC).

The principle behind this assay is that, after exonuclease V digestion, only double-stranded circular DNA survives in an acid-insoluble form. All DNA substrates and products that have free ends are digested to an acid-soluble form and are not retained on the filters. Therefore, only the 3 H-labeled attB linear DNA which ends up in circular form after both inter- and intramolecular integration is complete is resistant to digestion and is recovered as acid-insoluble product. Optimal enzyme and buffer formulations in the Clonase compositions therefore are those that give the highest levels of circularized 3 H-labeled attB-containing

sequences, as determined by highest cpm in the LSC. Although this assay was designed for optimization of GATEWAY™ BP Clonase™ Enzyme Mix compositions (Int + IHF), the same type of assay may be performed to optimize GATEWAY™ LR Clonase™ Enzyme Mix compositions (Int + IHF + Xis), except that the reaction mixtures would comprise 1000 ng of AttR (instead of AttP) and 600 ng of AttL (instead of AttB), and 40 ng of His₆-carboxy- tagged Xis (XisH6) in addition to the IntH6 and IHF.

Example 19: Testing Functionality of Entry and Destination Vectors

As part of assessment of the functionality of particular vectors of the invention, it is important to functionally test the ability of the vectors to recombine. This assessment can be carried out by performing a recombinational cloning reaction (as schematized in Figures 2, 4, and 5A and 5B, and as described herein and in commonly owned U.S. Application Nos. 08/486,139, filed June 7, 1995, 08/663,002, filed June 7, 1996 (now U.S. Patent No. 5,888,732), 09/005,476, filed January 12, 1998, and 09/177,387, filed October 23, 1998, the disclosures of all of which are incorporated by reference herein in their entireties), by transforming E. coli and scoring colony forming units. However, an alternative assay may also be performed to allow faster, more simple assessment of the functionality of a given Entry or Destination Vector by agarose gel electrophoresis. The following is a description of such an in vitro assay.

Materials and Methods:

Plasmid templates pEZC1301 (Figure 84) and pEZC1313 (Figure 85), each containing a single wild type att site, were used for the generation of PCR products containing attL or attR sites, respectively. Plasmid templates were linearized with A/*Not*I, phenol extracted, ethanol precipitated and dissolved in TE to a concentration of 1 ng/ μ l.

PCR primers (capital letters represent base changes from wildtype):

attL1 gggg agcct gctttttGtacAaa gttggcatta taaaaagca ttgc
attL2 gggg agcct gctttCttGtacAaa gttggcatta taaaaagca ttgc
attL right tgttgccggg aagctagagt aa

5

attR1 gggg Acaag ttTgtaCaaaaaagc tgaacgaga aacgtaaaat
attR2 gggg Acaag ttTgtaCaaGaaagc tgaacgaga aacgtaaaat
attR right ca gacggcatga tgaacctgaa

10 PCR primers were dissolved in TE to a concentration of 500 pmol/μl. Primer mixes were prepared, consisting of attL1 + attLright primers, attL2 + attLright primers, attR1 + attRright primers, and attR2 + attRright primers, each mix containing 20 pmol/μl of each primer.

PCR reactions:

15 1 μl plasmid template (1 ng)
 1 μl primer pairs (20 pmoles of each)
 3 μl of H₂O
 45 μl of Platinum PCR SuperMix® (Life Technologies, Inc.)

20 Cycling conditions (performed in MJ thermocycler):

 95°C/2 minutes

 94°C/30 seconds

 25 cycles of 58°C/30 seconds and 72°C/1.5 minutes

 72°C/5 minutes

25 5°C/hold

 The resulting attL PCR product was 1.5 kb, and the resulting attR PCR product was 1.0 kb.

30 PCR reactions were PEG/MgCl₂ precipitated by adding 150 μl H₂O and 100 μl of 3x PEG/ MgCl₂ solution followed by centrifugation. The PCR products were dissolved in 50 μl of TE. Quantification of the PCR product was performed by gel electrophoresis of 1 μl and was estimated to be 50-100 ng/μl.

Recombination reactions of PCR products containing attL or attR sites with GATEWAY™ plasmids was performed as follows:

8 µl of H₂O

2 µl of attL or attR PCR product (100-200 ng)

2 µl of GATEWAY™ plasmid (100 ng)

4 µl of 5x Destination buffer

4 µl of GATEWAY™ LR Clonase™ Enzyme Mix

20 µl total volume (the reactions can be scaled down to a 5 µl total volume by adjusting the volumes of the components to about ¼ of those shown above, while keeping the stoichiometries the same).

Clonase reactions were incubated at 25°C for 2 hours. 2 µl of proteinase K (2 mg/ml) was added to stop the reaction. 10 µl was then run on a 1 % agarose gel. Positive control reactions were performed by reacting attL1 PCR product (1.0 kb) with attR1 PCR product (1.5 kb) and by similarly reacting attL2 PCR product with attR2 PCR product to observe the formation of a larger (2.5 kb) recombination product. Negative controls were similarly performed by reacting attL1 PCR product with attR2 PCR product and vice versa or reactions of attL PCR product with an attL plasmid, etc.

In alternative assays, to test attB Entry vectors, plasmids containing single attP sites were used. Plasmids containing single att sites could also be used as recombination substrates in general to test all Entry and Destination vectors (*i.e.*, those containing attL, attR, attB and attP sites). This would eliminate the need to do PCR reactions.

Results:

Destination and Entry plasmids when reacted with appropriate att-containing PCR products formed linear recombinant molecules that could be easily visualized on an agarose gel when compared to control reactions containing no attL or attR PCR product. Thus, the functionality of Destination and Entry vectors constructed according to the invention may be determined either by carrying out the Destination or Entry recombination reactions as depicted in

Figures 2, 4, and 5A and 5B, or more rapidly by carrying out the linearization assay described in this Example.

Example 20: PCR Cloning Using Universal Adapter-Primers

As described herein, the cloning of PCR products using the GATEWAY™ PCR Cloning System (Life Technologies, Inc.; Rockville, MD) requires the addition of attB sites (attB1 and attB2) to the ends of gene-specific primers used in the PCR reaction. The protocols described in the preceding Examples suggest that the user add 29 bp (25 bp containing the attB site plus four G residues) to the gene-specific primer. It would be advantageous to high volume users of the GATEWAY™ PCR Cloning System to generate attB-containing PCR product using universal attB adapter-primers in combination with shorter gene-specific primers containing a specified overlap to the adapters. The following experiments demonstrate the utility of this strategy using universal attB adapter-primers and gene-specific primers containing overlaps of various lengths from 6 bp to 18 bp. The results demonstrate that gene-specific primers with overlaps of 10 bp to 18 bp can be used successfully in PCR amplifications with universal attB adapter-primers to generate full-length PCR products. These PCR products can then be successfully cloned with high fidelity in a specified orientation using the GATEWAY™ PCR Cloning System.

Methods and Results:

To demonstrate that universal attB adapter-primers can be used with gene-specific primers containing partial attB sites in PCR reactions to generate full-length PCR product, a small 256 bp region of the human hemoglobin cDNA was chosen as a target so that intermediate sized products could be distinguished from full-length products by agarose gel electrophoresis.

The following oligonucleotides were used:

B1-Hgb: GGGG ACA AGT TTG TAC AAA AAA GCA GGC T-5'-Hgb*
B2-Hgb: GGGG ACC ACT TTG TAC AAG AAA GCT GGG T-3'-Hgb**

-145-

18B1-Hgb: TG TAC AAA AAA GCA GGC T-5'-Hgb
 18B2-Hgb: TG TAC AAG AAA GCT GGG T-3'-Hgb
 15B1-Hgb: AC AAA AAA GCA GGC T-5'-Hgb
 15B2-Hgb: AC AAG AAA GCT GGG T-3'-Hgb
 5 12B1-Hgb: AA AAA GCA GGC T-5'-Hgb
 12B2-Hgb: AG AAA GCT GGG T-3'-Hgb
 11B1-Hgb: A AAA GCA GGC T-5'-Hgb
 11B2-Hgb: G AAA GCT GGG T-3'-Hgb
 10B1-Hgb: AAA GCA GGC T-5'-Hgb
 10 10B2-Hgb: AAA GCT GGG T-3'-Hgb
 9B1-Hgb: AA GCA GGC T-5'-Hgb
 9B2-Hgb: AA GCT GGG T-3'-Hgb
 8B1-Hgb: A GCA GGC T-5'-Hgb
 8B2-Hgb: A GCT GGG T-3'-Hgb
 15 7B1-Hgb: GCA GGC T-5'-Hgb
 7B2-Hgb: GCT GGG T-3'-Hgb
 6B1-Hgb: CA GGC T-5'-Hgb
 6B2-Hgb: CT GGG T-3'-Hgb

 20 attB1 adapter: GGGG ACA AGT TTG TAC AAA AAA GCA GGC T
 attB2 adapter: GGGG ACC ACT TTG TAC AAG AAA GCT GGG T

 * -5'-Hgb = GTC ACT AGC CTG TGG AGC AAG A
 ** -3'-Hgb = AGG ATG GCA GAG GGA GAC GAC A

25

30

The aim of these experiments was to develop a simple and efficient universal adapter PCR method to generate attB containing PCR products suitable for use in the GATEWAY™ PCR Cloning System. The reaction mixtures and thermocycling conditions should be simple and efficient so that the universal adapter PCR method could be routinely applicable to any PCR product cloning application.

35

PCR reaction conditions were initially found that could successfully amplify predominately full-length PCR product using gene-specific primers containing 18bp and 15 bp overlap with universal attB primers. These conditions are outlined below:

-146-

10 pmoles of gene-specific primers
10 pmoles of universal attB adapter-primers
1 ng of plasmid containing the human hemoglobin cDNA.
100 ng of human leukocyte cDNA library DNA.
5 μ l of 10x PLATINUM Taq HiFi® reaction buffer (Life Technologies, Inc.)
2 μ l of 50 mM MgSO₄
1 μ l of 10 mM dNTPs
0.2 μ l of PLATINUM Taq HiFi® (1.0 unit)
H₂O to 50 μ l total reaction volume

Cycling conditions:

25 x $\left\{ \begin{array}{l} 95^{\circ}\text{C}/5 \text{ min} \\ 94^{\circ}\text{C}/15 \text{ sec} \\ 50^{\circ}\text{C}/30 \text{ sec} \\ 68^{\circ}\text{C}/1 \text{ min} \\ 68^{\circ}\text{C}/5 \text{ min} \\ 5^{\circ}\text{C}/\text{hold} \end{array} \right.$

To assess the efficiency of the method, 2 μ l (1/25) of the 50 μ l PCR reaction was electrophoresed in a 3 % Agarose-1000 gel. With overlaps of 12 bp or less, smaller intermediate products containing one or no universal attB adapter predominated the reactions. Further optimization of PCR reaction conditions was obtained by titrating the amounts of gene-specific primers and universal attB adapter-primers. The PCR reactions were set up as outlined above except that the amounts of primers added were:

0, 1, 3 or 10 pmoles of gene-specific primers
0, 10, 30 or 100 pmoles of adapter-primers

Cycling conditions:

5 25 x

95°C/3 min
94°C/15 sec
50°C/45 sec
68°C/1 min
68°C/5 min
5°C/hold

10 The use of limiting amounts of gene-specific primers (3 pmoles) and excess adapter-primers (30 pmoles) reduced the amounts of smaller intermediate products. Using these reaction conditions the overlap necessary to obtain predominately full-length PCR product was reduced to 12 bp. The amounts of gene-specific and adapter-primers was further optimized in the following PCR reactions:

15 0, 1, 2 or 3 pmoles of gene-specific primers
0, 30, 40 or 50 pmoles of adapter-primers

Cycling conditions:

20 25 x

95°C/3 min
94°C/15 sec
48°C/1 min
68°C/1 min
68°C/5 min
5°C/hold

25 30 The use of 2 pmoles of gene-specific primers and 40 pmoles of adapter-primers further reduced the amounts of intermediate products and generated predominately full-length PCR products with gene-specific primers containing an 11 bp overlap. The success of the PCR reactions can be assessed in any PCR application by performing a no adapter control. The use of limiting amounts of gene-specific primers should give faint or barely visible bands when 1/25 to 1/10 of the PCR reaction is electrophoresed on a standard agarose gel. Addition of the

universal attB adapter-primers should generate a robust PCR reaction with a much higher overall yield of product.

PCR products from reactions using the 18 bp, 15 bp, 12 bp, 11 bp and 10 bp overlap gene-specific primers were purified using the CONCERT® Rapid PCR Purification System (PCR products greater than 500 bp can be PEG precipitated). The purified PCR products were subsequently cloned into an attP containing plasmid vector using the GATEWAY™ PCR Cloning System (Life Technologies, Inc.; Rockville, MD) and transformed into *E. coli*. Colonies were selected and counted on the appropriate antibiotic media and screened by PCR for correct inserts and orientation.

Raw PCR products (unpurified) from the attB adapter PCR of a plasmid clone of part of the human beta-globin (Hgb) gene were also used in GATEWAY™ PCR Cloning System reactions. PCR products generated with the full attB B1/B2-Hgb, the 12B1/B2, 11B1/B2 and 10B1/B2 attB overlap Hgb primers were successfully cloned into the GATEWAY™ pENTR21 attP vector (Figure 49). 24 colonies from each (24 x 4 = 96 total) were tested and each was verified by PCR to contain correct inserts. The cloning efficiency expressed as cfu/ml is shown below:

Primer Used	cfu/ml
Hgb full attB	8,700
Hgb 12 bp overlap	21,000
Hgb 11 bp overlap	20,500
Hgb 10 bp overlap	13,500
GFP control	1,300

Interestingly, the overlap PCR products cloned with higher efficiency than did the full attB PCR product. Presumably, and as verified by visualization on agarose gel, the adapter PCR products were slightly cleaner than was the full attB PCR product. The differences in colony output may also reflect the proportion of PCR product molecules with intact attB sites.

Using the attB adapter PCR method, PCR primers with 12 bp attB overlaps were used to amplify cDNAs of different sizes (ranging from 1 to 4 kb)

from a leukocyte cDNA library and from first strand cDNA prepared from HeLa total RNA. While three of the four cDNAs were able to be amplified by this method, a non-specific amplification product was also observed that under some conditions would interfere with the gene-specific amplification. This non-specific product was amplified in reactions containing the attB adapter-primers alone without any gene-specific overlap primers present. The non-specific amplification product was reduced by increasing the stringency of the PCR reaction and lowering the attB adapter PCR primer concentration.

These results indicate that the adapter-primer PCR approach described in this Example will work well for cloned genes. These results also demonstrate the development of a simple and efficient method to amplify PCR products that are compatible with the GATEWAY™ PCR Cloning System that allows the use of shorter gene-specific primers that partially overlap universal attB adapter-primers. In routine PCR cloning applications, the use of 12 bp overlaps is recommended. The methods described in this Example can thus reduce the length of gene-specific primers by up to 17 residues or more, resulting in a significant savings in oligonucleotide costs for high volume users of the GATEWAY™ PCR Cloning System. In addition, using the methods and assays described in this Example, one of ordinary skill can, using only routine experimentation, design and use analogous primer-adapters based on or containing other recombination sites or fragments thereof, such as attL, attR, attP, lox, FRT, etc.

Example 21: Mutational Analysis of the Bacteriophage Lambda attL and attR Sites: Determinants of att Site Specificity in Site-specific Recombination

To investigate the determinants of att site specificity, the bacteriophage lambda attL and attR sites were systematically mutagenized. As noted herein, the determinants of specificity have previously been localized to the 7 bp overlap region (TTTATAC, which is defined by the cut sites for the integrase protein and is the region where strand exchange takes place) within the 15 bp core region (GCTTTTTTATACTAA) which is identical in all four lambda att sites, attB, attP, attL and attR. This core region, however, has not heretofore been systematically

mutagenized and examined to define precisely which mutations produce unique changes in *att* site specificity.

Therefore, to examine the effect of *att* sequence on site specificity, mutant *attL* and *attR* sites were generated by PCR and tested in an *in vitro* site-specific recombination assay. In this way all possible single base pair changes within the 7 bp overlap region of the core *att* site were generated as well as five additional changes outside the 7 bp overlap but within the 15 bp core *att* site. Each *attL* PCR substrate was tested in the *in vitro* recombination assay with each of the *attR* PCR substrates.

Methods

To examine both the efficiency and specificity of recombination of mutant *attL* and *attR* sites, a simple *in vitro* site-specific recombination assay was developed. Since the core regions of *attL* and *attR* lie near the ends of these sites, it was possible to incorporate the desired nucleotide base changes within PCR primers and generate a series of PCR products containing mutant *attL* and *attR* sites. PCR products containing *attL* and *attR* sites were used as substrates in an *in vitro* reaction with GATEWAY™ LR Clonase™ Enzyme Mix (Life Technologies, Inc.; Rockville, MD). Recombination between a 1.5 kb *attL* PCR product and a 1.0 kb *attR* PCR product resulted in a 2.5 kb recombinant molecule that was monitored using agarose gel electrophoresis and ethidium bromide staining.

Plasmid templates pEZC1301 (Figure 84) and pEZC1313 (Figure 85), each containing a single wild type *attL* or *attR* site, respectively, were used for the generation of recombination substrates. The following list shows primers that were used in PCR reactions to generate the *attL* PCR products that were used as substrates in L x R Clonase reactions (capital letters represent changes from the wild-type sequence, and the underline represents the 7 bp overlap region within the 15 bp core *att* site; a similar set of PCR primers was used to prepare the *attR* PCR products containing matching mutations):

-151-

GATEWAY™ sites (note: attL2 sequence in GATEWAY™ plasmids begins "accca" while the attL2 site in this example begins "agcct" to reflect wild-type attL outside the core region.):

5

attL1: gggg agcct gcttttGtacAaa gttggcatta taaaaa-
agca ttgc

10

attL2: gggg agcct gcttttCttGtacAaa gttggcatta taaaaa-
agca ttgc

Wild-type:

15

attL0: gggg agcct gcttttttatactaa gttggcatta taaaaa-
agca ttgc

Single base changes from wild-type:

20

attLT1A: gggg agcct gcttttAttatactaa gttggcatta taaaaa-
agca ttgc

attLT1C: gggg agcct gcttttCttatactaa gttggcatta taaaaa-
agca ttgc

25

attLT1G: gggg agcct gcttttGttatactaa gttggcatta taaaaa-
agca ttgc

attLT2A: gggg agcct gcttttAtatactaa gttggcatta taaaaa-
agca ttgc

30

attLT2C: gggg agcct gcttttCtatactaa gttggcatta taaaaa-
agca ttgc

attLT2G: gggg agcct gcttttGtatactaa gttggcatta taaaaa-
aagca ttgc

35

-152-

attLT3A: gggg agcct gcttttttAataactaa gttggcatta taaaa-
aagca ttgc

5 attLT3C: gggg agcct gcttttttCataactaa gttggcatta taaaa-
aagca ttgc

10 attLT3G: gggg agcct gcttttttGataactaa gttggcatta taaaa-
aagca ttgc

15 attLA4C: gggg agcct gcttttttCtaactaa gttggcatta taaaa-
aagca ttgc

attLA4G: gggg agcct gcttttttGtaactaa gttggcatta taaaa-
aagca ttgc

20 attLA4T: gggg agcct gcttttttTtaactaa gttggcatta taaaa-
aagca ttgc

25 attLT5A: gggg agcct gcttttttAaactaa gttggcatta taaaa-
aagca ttgc

attLT5C: gggg agcct gcttttttCaactaa gttggcatta taaaa-
aagca ttgc

30 attLT5G: gggg agcct gcttttttGaactaa gttggcatta taaaa-
aagca ttgc

35 attLA6C: gggg agcct gcttttttatCctaa gttggcatta taaaa-
aagca ttgc

-153-

attLA6G: gggg agcct gcttttttatGctaa gttggcatta taaaa-
aagca ttgc

5 attLA6T: gggg agcct gcttttttatTctaa gttggcatta taaaa-
aagca ttgc

10 attLC7A: gggg agcct gcttttttataAtaa gttggcatta taaaa-
aagca ttgc

15 attLC7G: gggg agcct gcttttttataGtaa gttggcatta taaaa-
aagca ttgc

attLC7T: gggg agcct gcttttttataTtaa gttggcatta taaaa-
aagca ttgc

Single base changes outside of the 7 bp overlap:

20 attL8: gggg agcct Acttttttatatactaa gttggcatta taaaa-
aagca ttgc

25 attL9: gggg agcct gcCtttttatatactaa gttggcatta taaaaa-
agca ttgc

attL10: gggg agcct gcttCtttatatactaa gttggcatta taaaaa-
agca ttgc

30 attL14: gggg agcct gcttttttatacCaa gttggcatta taaaaa-
agca ttgc

35 attL15: gggg agcct gcttttttatatactaG gttggcatta taaaaa-
agca ttgc

Note: additional vectors wherein the first nine bases are gggg agcca (*i.e.*, substituting an adenine for the thymine in the position immediately preceding the 15-bp core region), which may or may not contain the single base pair substitutions (or deletions) outlined above, can also be used in these experiments.

Recombination reactions of *attL*- and *attR*-containing PCR products was performed as follows:

8 µl of H₂O
2 µl of *attL* PCR product (100 ng)
2 µl of *attR* PCR product (100 ng)
4 µl of 5x buffer
4 µl of GATEWAY™ LR Clonase™ Enzyme Mix
20 µl total volume

Clonase reactions were incubated at 25°C for 2 hours.

2 µl of 10X Clonase stop solution (proteinase K, 2 mg/ml) were added to stop the reaction.

10 µl were run on a 1 % agarose gel.

Results

Each *attL* PCR substrate was tested in the *in vitro* recombination assay with each of the *attR* PCR substrates. Changes within the first three positions of the 7 bp overlap (TTTATAC) strongly altered the specificity of recombination. These mutant *att* sites each recombined as well as the wild-type, but only with their cognate partner mutant; they did not recombine detectably with any other *att* site mutant. In contrast, changes in the last four positions (TTTATAC) only partially altered specificity; these mutants recombined with their cognate mutant as well as wild-type *att* sites and recombined partially with all other mutant *att* sites except for those having mutations in the first three positions of the 7 bp

overlap. Changes outside of the 7 bp overlap were found not to affect specificity of recombination, but some did influence the efficiency of recombination.

Based on these results, the following rules for *att* site specificity were determined:

- Only changes within the 7 bp overlap affect specificity.
- Changes within the first 3 positions strongly affect specificity.
- Changes within the last 4 positions weakly affect specificity.

Mutations that affected the overall efficiency of the recombination reaction were also assessed by this method. In these experiments, a slightly increased (less than 2-fold) recombination efficiency with *attLT1A* and *attLC7T* substrates was observed when these substrates were reacted with their cognate *attR* partners. Also observed were mutations that decreased recombination efficiency (approximately 2-3 fold), including *attLA6G*, *attL14* and *attL15*. These mutations presumably reflect changes that affect Int protein binding at the core *att* site.

The results of these experiments demonstrate that changes within the first three positions of the 7 bp overlap (TTTATAC) strongly altered the specificity of recombination (*i.e.*, *att* sequences with one or more mutations in the first three thymidines would only recombine with their cognate partners and would not cross-react with any other *att* site mutation). In contrast, mutations in the last four positions (TTTATAC) only partially altered specificity (*i.e.*, *att* sequences with one or more mutations in the last four base positions would cross-react partially with the wild-type *att* site and all other mutant *att* sites, except for those having mutations in one or more of the first three positions of the 7 bp overlap). Mutations outside of the 7 bp overlap were not found to affect specificity of recombination, but some were found to influence (*i.e.*, to cause a decrease in) the efficiency of recombination.

Example 22: Discovery of Att Site Mutations That Increase the Cloning Efficiency of GATEWAY™ Cloning Reactions

In experiments designed to understand the determinants of *att* site specificity, point mutations in the core region of *attL* were made. Nucleic acid molecules containing these mutated *attL* sequences were then reacted in an LR

reaction with nucleic acid molecules containing the cognate *attR* site (*i.e.*, an *attR* site containing a mutation corresponding to that in the *attL* site), and recombinational efficiency was determined as described above. Several mutations located in the core region of the *att* site were noted that either slightly increased (less than 2-fold) or decreased (between 2-4-fold) the efficiency of the recombination reaction (Table 3).

Table 3. *Effects of attL mutations on Recombination Reactions.*

<u>Site</u>	<u>Sequence</u>	<u>Effect on Recombination</u>
attL0	agcctgcttttttataactaagttggcatta	
attL5	agcctgctttAttataactaagttggcatta	slightly increased
attL6	agcctgctttttttataTtaagttggcatta	slightly increased
attL13	agcctgctttttttatGctaagttggcatta	decreased
attL14	agcctgctttttttatacCaagttggcatta	decreased
attL15	agcctgctttttttatactaGgttggcatta	decreased
consensus	CAACTTnnTnnnAnnAAGTTG	

It was also noted that these mutations presumably reflected changes that either increased or decreased, respectively, the relative affinity of the integrase protein for binding the core *att* site. A consensus sequence for an integrase core-binding site (CAACTTNNT) has been inferred in the literature but not directly tested (see, *e.g.*, Ross and Landy, *Cell* 33:261-272 (1983)). This consensus core integrase-binding sequence was established by comparing the sequences of each of the four core *att* sites found in *attP* and *attB* as well as the sequences of five non-*att* sites that resemble the core sequence and to which integrase has been shown to bind in vitro. These experiments suggest that many more *att* site mutations might be identified which increase the binding of integrase to the core *att* site and thus increase the efficiency of GATEWAY™ cloning reactions.

Example 23: Effects of Core Region Mutations on Recombination Efficiency

To directly compare the cloning efficiency of mutations in the att site core region, single base changes were made in the attB2 site of an attB1-TET-attB2 PCR product. Nucleic acid molecules containing these mutated attB2 sequences were then reacted in a BP reaction with nucleic acid molecules containing non-cognate attP sites (*i.e.*, wildtype attP2), and recombinational efficiency was determined as described above. The cloning efficiency of these mutant attB2 containing PCR products compared to standard attB1-TET-attB2 PCR product are shown in Table 4.

Table 4. Efficiency of Recombination With Mutated attB2 Sites.

<u>Site</u>	<u>Sequence</u>	<u>Mutation</u>	<u>Cloning Efficiency</u>
attB0	tcaagttagtagtataaaaaagcaggct		
attB1	ggggacaagttgtacaaaaagcaggct		
attB2	ggggaccactttgtacaagaaagctgggt		100%
attB2.1	gggggAacactttgtacaagaaagctgggt	C→A	40%
attB2.2	ggggacAactttgtacaagaaagctgggt	C→A	131%
attB2.3	ggggaccCctttgtacaagaaagctgggt	A→C	4%
attB2.4	ggggaccaAttgtacaagaaagctgggt	C→A	11%
attB2.5	ggggaccacGttgtacaagaaagctgggt	T→G	4%
attB2.6	ggggaccactGtgtacaagaaagctgggt	T→G	6%
attB2.7	ggggaccacttGgtacaagaaagctgggt	T→G	1%
attB2.8	ggggaccactttTtacaagaaagctgggt	G→T	0.5%

As noted above, a single base change in the attB2.2 site increased the cloning efficiency of the attB1-TET-attB2.2 PCR product to 131% compared to the attB1-TET-attB2 PCR product. Interestingly, this mutation changes the integrase core binding site of attB2 to a sequence that matches more closely the proposed consensus sequence.

Additional experiments were performed to directly compare the cloning efficiency of an attB1-TET-attB2 PCR product with a PCR product that contained attB sites containing the proposed consensus sequence (*see* Example 22) of an integrase core binding site. The following attB sites were used to amplify attB-TET PCR products:

attB1 ggggacaagtttgtacaaaaaagcaggct
attB1.6 ggggacaaCtttgtacaaaaaagTTggct
attB2 ggggaccactttgtacaagaaagctgggt
attB2.10 ggggacAactttgtacaagaaagTtgggt

BP reactions were carried out between 300 ng (100 fmoles) of pDONR201 (Figure 49A) with 80 ng (80 fmoles) of attB-TET PCR product in a 20 µl volume with incubation for 1.5 hrs at 25°C, creating pENTR201-TET Entry clones. A comparison of the cloning efficiencies of the above-noted attB sites in BP reactions is shown in Table 5.

Table 5. Cloning efficiency of BP Reactions.

PCR product	CFU/ml	Fold Increase
B1-tet-B2	7,500	
B1.6-tet-B2	12,000	1.6 x
B1-tet-B2.10	20,900	2.8 x
B1.6-tet-B2.10	30,100	4.0 x

These results demonstrate that attB PCR products containing sequences that perfectly match the proposed consensus sequence for integrase core binding sites can produce Entry clones with four-fold higher efficiency than standard Gateway attB1 and attB2 PCR products.

The entry clones produced above were then transferred to pDEST20 (Figure 40A) via LR reactions (300 ng (64 fmoles) pDEST20 mixed with 50 ng (77 fmoles) of the respective pENTR201-TET Entry clone in 20 µl volume; incubated for 1 hr incubation at 25°C). The efficiencies of cloning for these reactions are compared in Table 6.

Table 6. Cloning Efficiency of LR Reactions.

pENTR201-TET x pDEST20	CFU/ml	Fold Increase
L1-tet-L2	5,800	
L1.6-tet-L2	8,000	1.4
L1-tet-L2.10	10,000	1.7
L1.6-tet-L2.10	9,300	1.6

These results demonstrate that the mutations introduced into attB1.6 and attB2.10 that transfer with the gene into entry clones slightly increase the efficiency of LR reactions. Thus, the present invention encompasses not only mutations in *attB* sites that increase recombination efficiency, but also to the corresponding mutations that result in the *attL* sites created by the BP reaction.

To examine the increased cloning efficiency of the attB1.6-TET-attB2.10 PCR product over a range of PCR product amounts, experiments analogous to those described above were performed in which the amount of attB PCR product was titrated into the reaction mixture. The results are shown in Table 7.

Table 7. Titration of attB PCR products.

Amount of attB PCR product (ng)	PCR product	CFU/ml	Fold Increase
20	attB1-TET-attB2	3,500	6.1
	attB1.6-TET-attB2.10	21,500	
50	attB1-TET-attB2	9,800	5.0
	attB1.6-TET-attB2.10	49,000	
100	attB1-TET-attB2	18,800	2.8
	attB1.6-TET-attB2.10	53,000	
200	attB1-TET-attB2	19,000	2.5
	attB1.6-TET-attB2.10	48,000	

These results demonstrate that as much as a six-fold increase in cloning efficiency is achieved with the attB1.6-TET-attB2.10 PCR product as compared to the standard attB1-TET-attB2 PCR product at the 20 ng amount.

Example 24: Determination of attB Sequence Requirements for Optimum Recombination Efficiency

To examine the sequence requirements for attB and to determine which attB sites would clone with the highest efficiency from populations of degenerate attB sites, a series of experiments was performed. Degenerate PCR primers were designed which contained five bases of degeneracy in the B-arm of the attB site. These degenerate sequences would thus transfer with the gene into Entry clone in BP reactions and subsequently be transferred with the gene into expression clones in LR reactions. The populations of degenerate attB and attL sites could thus be cycled from attB to attL back and forth for any number of cycles. By altering the reaction conditions at each transfer step (for example by decreasing the reaction time and/or decreasing the concentration of DNA) the reaction can be made increasingly more stringent at each cycle and thus enrich for populations of attB and attL sites that react more efficiently.

The following degenerate PCR primers were used to amplify a 500 bp fragment from pUC18 which contained the lacZ alpha fragment (only the attB portion of each primer is shown):

```

attB1      GGGG ACAAGTTTGTACAAA AAAGC AGGCT
attB1n16-20 GGGG ACAAGTTTGTACAAA nnnnn AGGCT
attB1n21-25 GGGG ACAAGTTTGTACAAA AAAGC nnnnn

attB2      GGGG ACCACTTTGTACAAG AAAGC TGGGT
attB2n16-20 GGGG ACCACTTTGTACAAG nnnnn TGGGT
attB2n21-25 GGGG ACCACTTTGTACAAG AAAGC nnnnn

```

The starting population size of degenerate att sites is 4^5 or 1024 molecules. Four different populations were transferred through two BP reactions and two LR reactions. Following transformation of each reaction, the population of transformants was amplified by growth in liquid media containing the appropriate selection antibiotic. DNA was prepared from the population of clones by alkaline

-161-

lysis miniprep and used in the next reaction. The results of the BP and LR cloning reactions are shown below.

BP-1, overnight reactions

	cfu/ml	percent of control
attB1-LacZa-attB2	78,500	100 %
attB1n16-20-LacZa-attB2	1,140	1.5 %
attB1n21-25-LacZa-attB2	11,100	14 %
attB1-LacZa-attB2n16-20	710	0.9 %
attB1-LacZa-attB2n21-25	16,600	21 %

LR-1, pENTR201-LacZa x pDEST20/*Eco*RI, 1hr reactions

	cfu/ml	percent of control
attL1-LacZa-attL2	20,000	100 %
attL1n16-20-LacZa-attL2	2,125	11 %
attL1n21-25-LacZa-attL2	2,920	15 %
attL1-LacZa-attL2n16-20	3,190	16 %
attL1-LacZa-attL2n21-25	1,405	7 %

BP-2, pEXP20-LacZa/*Sca*I x pDONR 201, 1hr reactions

	cfu/ml	percent of control
attB1-LacZa-attB2	48,600	100 %
attB1n16-20-LacZa-attB2	22,800	47 %
attB1n21-25-LacZa-attB2	31,500	65 %
attB1-LacZa-attB2n16-20	42,400	87 %
attB1-LacZa-attB2n21-25	34,500	71 %

LR-2, pENTR201-LacZa x pDEST6/*Nco*I, 1hr reactions

	cfu/ml	percent of control
attL1-LacZa-attL2	23,000	100 %
attL1n16-20-LacZa-attL2	49,000	213 %
attL1n21-25-LacZa-attL2	18,000	80 %
attL1-LacZa-attL2n16-20	37,000	160 %
attL1-LacZa-attL2n21-25	57,000	250 %

These results demonstrate that at each successive transfer, the cloning efficiency of the entire population of att sites increases, and that there is a great deal of flexibility in the definition of an attB site. Specific clones may be isolated from the above reactions, tested individually for recombination efficiency, and

sequenced. Such new specificities may then be compared to known examples to guide the design of new sequences with new recombination specificities. In addition, based on the enrichment and screening protocols described herein, one of ordinary skill can easily identify and use sequences in other recombination sites, *e.g.*, other *att* sites, *lox*, FRT, etc., that result in increased specificity in the recombination reactions using nucleic acid molecules containing such sequences.

Example 25: Design of att Site PCR Adapter-Primers

Additional studies were performed to design gene-specific primers with 12bp of attB1 and attB2 at their 5'-ends. The optimal primer design for *att*-containing primers is the same as for any PCR primers: the gene-specific portion of the primers should ideally have a T_m of $> 50^\circ\text{C}$ at 50 mM salt (calculation of T_m is based on the formula $59.9 + 41(\%GC) - 675/n$).

Primers:

12bp attB1: AA AAA GCA GGC TNN - forward gene-specific primer

12bp attB2: A GAA AGC TGG GTN - reverse gene-specific primer

attB1 adapter primer: GGGGACAAGTTTGTACAAAAAAGCAGGCT

attB2 adapter primer: GGGGACCACTTTGTACAAGAAAGCTGGGT

Protocol:

(1) Mix 200 ng of cDNA library or 1 ng of plasmid clone DNA (alternatively, genomic DNA or RNA could be used) with 10 pmoles of gene specific primers in a 50 μl PCR reaction, using one or more polypeptides having DNA polymerase activity such as those described herein. (The addition of greater than 10 pmoles of gene-specific primers can decrease the yield of attB PCR product. In addition, if RNA is used, a standard reverse transcriptase-PCR (RT-

PCR) protocol should be followed; *see, e.g.,* Gerard, G.F., *et al.*, *FOCUS* 11:60 (1989); Myers, T.W., and Gelfand, D.H., *Biochem.* 30:7661 (1991); Freeman, W.N., *et al.*, *BioTechniques* 20:782 (1996); and U.S. Application No. 09/064,057, filed April 22, 1998, the disclosures of all of which are incorporated herein by reference.)

1st PCR profile:

(a) 95°C for 3 minutes

(b) 10 cycles of:

(i) 94°C for 15 seconds

(ii) 50°C* for 30 seconds

(iii) 68°C for 1 minute/kb of target amplicon

(c) 68°C for 5 minutes

(d) 10°C hold

*The optimal annealing temperature is determined by the calculated T_m of the gene-specific part of the primer.

(2) Transfer 10 µl to a 40 µl PCR reaction mix containing 35 pmoles each of the attB1 and attB2 adapter primers.

2nd PCR profile:

(a) 95°C for 1 minute

(b) 5 cycles of:

(i) 94°C for 15 seconds

(ii) 45°C* for 30 seconds

(iii) 68°C for 1 minute/kb of target amplicon

(c) 15-20 cycles** of:

(i) 94°C for 15 seconds

(ii) 55°C* for 30 seconds

(iii) 68°C for 1 minute/kb of target amplicon

(d) 68°C for 5 minutes

(e) 10°C hold

5 *The optimal annealing temperature is determined by the calculated T_m of the gene-specific part of the primer.

 **15 cycles is sufficient for low complexity targets.

Notes:

- 10 1. It is useful to perform a no-adaptor primer control to assess the yield of attB PCR product produced.
2. Linearized template usually results in slightly greater yield of PCR product.

15

Example 26: One-Tube Recombinational Cloning Using the GATEWAY™ Cloning System

20 To provide for easier and more rapid cloning using the GATEWAY™ cloning system, we have designed a protocol whereby the BP and LR reactions may be performed in a single tube (a "one-tube" protocol). The following is an example of such a one-tube protocol; in this example, an aliquot of the BP reaction is taken before adding the LR components, but the BP and LR reactions may be performed in a one-tube protocol without first taking the BP aliquot:

25

<u>Reaction Component</u>	<u>Volume</u>
attB DNA (100-200 ng/25 µl reaction)	1-12.5 µl
attP DNA (pDONR201) 150 ng/µl	2.5 µl
5X BP Reaction Buffer	5.0 µl
Tris-EDTA	(to 20 µl)
<u>BP Clonase</u>	<u>5.0 µl</u>
Total vol.	25 µl

30

-165-

After the above components were mixed in a single tube, the reaction mixtures were incubated for 4 hours at 25°C. A 5 µl aliquot of reaction mixture was removed, and 0.5 µl of 10X stop solution was added to this reaction mixture and incubated for 10 minutes at 37°C. Competent cells were then transformed with 1-2 µl of the BP reaction per 100 µl of cells; this transformation yielded colonies of Entry Clones for isolation of individual Entry Clones and for quantitation of the BP Reaction efficiency.

To the remaining 20 µl of BP reaction mixture, the following components of the LR reaction were added:

<u>Reaction Component</u>	<u>Final Concentration</u>	<u>Volume Added</u>
NaCl	0.75 M	1 µl
Destination Vector	150 ng/ul	3 µl
<u>LR Clonase</u>		<u>6 µl</u>
Total vol.		30 µl

After the above components were mixed in a single tube, the reaction mixtures were incubated for 2 hours at 25°C. 3 µl of 10X stop solution was added, and the mixture was incubated for 10 minutes at 37°C. Competent cells were then transformed with 1-2 µl of the reaction mixture per 100 µl of cells

Notes:

1. If desired, the Destination Vector can be added to the initial BP reaction.
2. The reactions can be scaled down by 2x, if desired.
3. Shorter incubation times for the BP and/or LR reactions can be used (scaled to the desired cloning efficiencies of the reaction), but a lower number of colonies will typically result.
4. To increase the number of colonies obtained by several fold, incubate the BP reaction for 6-20 hours and increase the LR reaction to 3 hours. Electroporation also works well with 1-2 ul of the PK-treated reaction mixture.

5. PCR products greater than about 5 kb may show significantly lower cloning efficiency in the BP reaction. In this case, we recommend using a one-tube reaction with longer incubation times (e.g., 6-18 hours) for both the BP and LR steps.

Example 27: Relaxation of Destination Vectors During the LR Reaction

To further optimize the LR Reaction, the composition of the LR Reaction buffer was modified from that described above and this modified buffer was used in a protocol to examine the impact of enzymatic relaxation of Destination Vectors during the LR Reaction.

LR Reactions were set up as usual (see, e.g., Example 6), except that 5X BP Reaction Buffer (see Example 5) was used for the LR Reaction. To accomplish Destination Vector relaxation during the LR Reaction, Topoisomerase I (Life Technologies, Inc., Rockville, MD; Catalogue No. 38042-016) was added to the reaction mixture at a final concentration of ~15U per µg of total DNA in the reaction (for example, for reaction mixtures with a total of 400ng DNA in the 20 µl LR Reaction, ~6units of Topoisomerase I was added). Reaction mixtures were set up as follows:

<u>Reaction Component</u>	<u>Volume</u>
ddH ₂ O	6.5 µl
4X BP Reaction Buffer	5 µl
100ng single chain/linear pENTR CAT, 50 ng/µl	2 µl
300ng single chain/linear pDEST6, 150ng/µl	2 µl
Topoisomerase I, 15 U/ml	0.5 µl
LR Clonase	4 µl

Reaction mixtures were incubated at 25°C for 1hour, and 2 µl of 2 µg/µl Proteinase K was then added and mixtures incubated for 10 minutes at 37°C to stop the LR Reaction. Competent cells were then transformed as described in the preceding examples. The results of these studies demonstrated that relaxation of

substrates in the LR reaction using Topoisomerase I resulted in a 2- to 10-fold increase in colony output compared to those LR reactions performed without including Topoisomerase I.

5 Having now fully described the present invention in some detail by way of illustration and example for purposes of clarity of understanding, it will be obvious to one of ordinary skill in the art that the same can be performed by modifying or changing the invention within a wide and equivalent range of conditions, formulations and other parameters without affecting the scope of the invention or
10 any specific embodiment thereof, and that such modifications or changes are intended to be encompassed within the scope of the appended claims.

 All publications, patents and patent applications mentioned in this specification are indicative of the level of skill of those skilled in the art to which this invention pertains, and are herein incorporated by reference to the same extent
15 as if each individual publication, patent or patent application was specifically and individually indicated to be incorporated by reference.

167.1

Applicant's or agent's file reference number	0942.~J8PC03	International application No. PCT/US	00/05432
---	--------------	---	-----------------

INDICATIONS RELATING TO DEPOSITED MICROORGANISM
OR OTHER BIOLOGICAL MATERIAL
(PCT Rule 13bis)

REC'D 17 APR 2000

WIPO PCT

A. The indications made below relate to the microorganism referred to in the description on page 52, line 31.

B. IDENTIFICATION OF DEPOSIT

Further deposits are identified on an additional sheet ☒

Name of depositary institution
Agricultural Research Culture Collection (NRRL)
International Depository Authority

Address of depositary institution (including postal code and country)

1815 N. University Street
Peoria, Illinois 61604
United States of America

Date of deposit
February 27, 1999

Accession Number
NRRL B-30099

C. ADDITIONAL INDICATIONS (leave blank if not applicable)

This information is continued on an additional sheet ☐

Escherichia coli DB3.1(pAHPKan) or Escherichia coli DB3.1(pAtHPKan)

D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)

E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)

The indications listed below will be submitted to the international Bureau later (specify the general nature of the indications, e.g., "Accession Number of Deposit")

For receiving Office use only

☒ This sheet was received with the international application

Authorized officer

Signature: [Signature]
Name: [Name]
Address: [Address]

For International Bureau use only

☐ This sheet was received by the International Bureau on:

Authorized officer

167.2

Applicant's or agent's file reference number	0942.468PC03	International application No. tl PCT/US 00/05432
---	--------------	--

**INDICATIONS RELATING TO DEPOSITED MICROORGANISM
OR OTHER BIOLOGICAL MATERIAL**
(PCT Rule 13bis)

REC'D 17 APR 2000

WIPO PCT

A. The indications made below relate to the microorganism referred to in the description on page <u>55</u> , line <u>16</u> .	
B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input checked="" type="checkbox"/>	
Name of depositary institution Agricultural Research Culture Collection (NRRL) International Depositary Authority	
Address of depositary institution (including postal code and country) 1815 N. University Street Peoria, Illinois 61604 United States of America	
Date of deposit February 27, 1999	Accession Number NRRL B-30100
C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is continued on an additional sheet <input type="checkbox"/>	
Escherichia coli DB3.1(pENTR-1A)	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)	
E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)	
The indications listed below will be submitted to the international Bureau later (specify the general nature of the indications, e.g., "Accession Number of Deposit")	

For receiving Office use only	For International Bureau use only
<input checked="" type="checkbox"/> This sheet was received with the international application	<input type="checkbox"/> This sheet was received by the International Bureau on:
Authorized officer Barbara Fridis PCT/US00/05432	Authorized officer

Applicant's or agent's file reference number	0942.468PC03	International application No. tb- 00/05432
--	--------------	--

**INDICATIONS RELATING TO DEPOSITED MICROORGANISM
OR OTHER BIOLOGICAL MATERIAL**
(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page 55, line 16.

B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet ☒

Name of depositary institution
Agricultural Research Culture Collection (NRRL)
International Depository Authority

Address of depositary institution (including postal code and country)

1815 N. University Street
Peoria, Illinois 61604
United States of America

Date of deposit February 27, 1999	Accession Number NRRL B-30101
--------------------------------------	----------------------------------

C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is continued on an additional sheet ☐

Escherichia coli DB3.1(pENTR-2B)

D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE <i>(if the indications are not for all designated States)</i>	

E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)

The indications listed below will be submitted to the international Bureau later (*specify the general nature of the indications, e.g., "Accession Number of Deposit"*)

For receiving Office use only	For International Bureau use only
<input checked="" type="checkbox"/> This sheet was received with the international application	<input type="checkbox"/> This sheet was received by the International Bureau on:
Authorized officer <i>Walter Fricke</i> <i>87</i> ST. Gallen, Switzerland 1987-05-15 15:23:00 (E)	Authorized officer

167.4

Applicant's or agent's file reference number	0942.468PC03	International Application No. PCT/US 00/05432
---	--------------	--

INDICATIONS RELATING TO DEPOSITED MICROORGANISM
OR OTHER BIOLOGICAL MATERIAL
(PCT Rule 13bis)

REC'D 17 APR 2000

WIPO PCT

A. The indications made below relate to the microorganism referred to in the description on page 55, line
16.

B. IDENTIFICATION OF DEPOSIT

Further deposits are identified on an additional sheet ☒

Name of depositary institution
Agricultural Research Culture Collection (NRRL)
International Depository Authority

Address of depositary institution (including postal code and country)

1815 N. University Street
Peoria, Illinois 61604
United States of America

Date of deposit
February 27, 1999

Accession Number
NRRL B-30102

C. ADDITIONAL INDICATIONS (leave blank if not applicable)

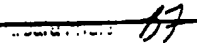
This information is continued on an additional sheet ☐

Escherichia coli DB3.1(pENTR-3C)

D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)

E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)

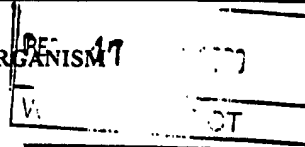
The indications listed below will be submitted to the international Bureau later (specify the general nature of the indications, e.g.,
"Accession Number of Deposit")

For receiving Office use only	For International Bureau use only
<input checked="" type="checkbox"/> This sheet was received with the international application	<input type="checkbox"/> This sheet was received by the International Bureau on:
Authorized officer  100-3000 (F)	Authorized officer

167.5

Applicant's or agent's file reference number	0942.468PC03	International application No. tb PCT/US 00/05432
---	--------------	---

**INDICATIONS RELATING TO DEPOSITED MICROORGANISM
OR OTHER BIOLOGICAL MATERIAL**
(PCT Rule 13bis)



A. The indications made below relate to the microorganism referred to in the description on page 8	REC-17 APR 2000 WIPO PCT
--	-----------------------------

B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet ☒

Name of depositary institution
Agricultural Research Culture Collection (NRRL)
International Depository Authority

Address of depositary institution (including postal code and country)

1815 N. University Street
Peoria, Illinois 61604
United States of America

Date of deposit
February 27, 1999

Accession Number
NRRL B-30103

C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is continued on an additional sheet ☐

Escherichia coli DB3.1(pEZC15101)

D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)

E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)

The indications listed below will be submitted to the international Bureau later (specify the general nature of the indications, e.g., "Accession Number of Deposit")

For receiving Office use only	For International Bureau use only
<input checked="" type="checkbox"/> This sheet was received with the international application	<input type="checkbox"/> This sheet was received by the International Bureau on:
Authorized officer [Signature] [Stamp: RECEIVED, APR 17 2000, WIPO]	Authorized officer

167.6

Applicant's or agent's file reference number	0942.468PC03	International application No. 1.	PCT/US 00/05432
--	--------------	----------------------------------	-----------------

**INDICATIONS RELATING TO DEPOSITED MICROORGANISM
OR OTHER BIOLOGICAL MATERIAL**
(PCT Rule 13bis)

REC'D 17

WIPO

A. The indications made below relate to the microorganism referred to in the description on page 54, line 9.

B. IDENTIFICATION OF DEPOSITFurther deposits are identified on an additional sheet ☒

Name of depositary institution
Agricultural Research Culture Collection (NRRL)
International Depositary Authority

Address of depositary institution (including postal code and country)

1815 N. University Street
Peoria, Illinois 61604
United States of America

Date of deposit
February 27, 1999

Accession Number
NRRL B-30104

C. ADDITIONAL INDICATIONS (leave blank if not applicable)This information is continued on an additional sheet ☐

Escherichia coli DB3.1(pEZC15102)

D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)**E. SEPARATE FURNISHING OF INDICATIONS** (leave blank if not applicable)

The indications listed below will be submitted to the international Bureau later (specify the general nature of the indications, e.g., "Accession Number of Deposit")

For receiving Office use only

☒ This sheet was received with the international application

Authorized officer

Barbara F. H. H. H.

10/22/99

10/22/99

10/22/99

For International Bureau use only

☐ This sheet was received by the International Bureau on:

Authorized officer

167.7

Applicant's or agent's file reference number	0942.468PC03	International application No. PCT/US	00/05432
---	--------------	---	-----------------

INDICATIONS RELATING TO DEPOSITED MICROORGANISMS
OR OTHER BIOLOGICAL MATERIAL
(PCT Rule 13bis)

RECEIVED 17 APR 2000

V T

A. The indications made below relate to the microorganism referred to in the description on page 54, line 9.

B. IDENTIFICATION OF DEPOSITFurther deposits are identified on an additional sheet ☒

Name of depositary institution
Agricultural Research Culture Collection (NRRL)
International Depositary Authority

Address of depositary institution (including postal code and country)

1815 N. University Street
Peoria, Illinois 61604
United States of America

Date of deposit
February 27, 1999

Accession Number
NRRL B-30105

C. ADDITIONAL INDICATIONS (leave blank if not applicable)This information is continued on an additional sheet ☐

Escherichia coli DB3.1(pEZC15103)

D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)**E. SEPARATE FURNISHING OF INDICATIONS** (leave blank if not applicable)

The indications listed below will be submitted to the international Bureau later (specify the general nature of the indications, e.g., "Accession Number of Deposit")

For receiving Office use only

<input checked="" type="checkbox"/> This sheet was received with the international application
Authorized officer: <i>Richard E. Hickey</i>

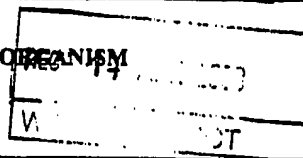
For International Bureau use only

<input type="checkbox"/> This sheet was received by the International Bureau on:
Authorized officer:

167.8

Applicant's or agent's file reference number	0942.408PC03	International application No. of PCT/US 00/05432
---	--------------	--

INDICATIONS RELATING TO DEPOSITED MICROORGANISM
OR OTHER BIOLOGICAL MATERIAL
(PCT Rule 13bis)



A. The indications made below relate to the microorganism referred to in the description on page 51, line 20-21.

B. IDENTIFICATION OF DEPOSIT

Further deposits are identified on an additional sheet ☒

Name of depositary institution
Agricultural Research Culture Collection (NRRL)
International Depository Authority

Address of depositary institution (including postal code and country)

1815 N. University Street
Peoria, Illinois 61604
United States of America

Date of deposit
February 27, 1999

Accession Number
NRRL B-30108

C. ADDITIONAL INDICATIONS (leave blank if not applicable)

This information is continued on an additional sheet ☐

Escherichia coli DB10B(pCMVSPORT6)

D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)

E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)

The indications listed below will be submitted to the international Bureau later (specify the general nature of the indications, e.g., "Accession Number of Deposit")

For receiving Office use only	For International Bureau use only
<input checked="" type="checkbox"/> This sheet was received with the international application	<input type="checkbox"/> This sheet was received by the International Bureau on:
Authorized officer Barbara Fricke <i>BF</i>	Authorized officer

WHAT IS CLAIMED IS:

1. An isolated nucleic acid molecule comprising a nucleotide sequence selected from the group of nucleotide sequences consisting of an attB1 nucleotide sequence as set forth in Figure 9, an attB2 nucleotide sequence as set forth in Figure 9, an attP1 nucleotide sequence as set forth in Figure 9, an attP2 nucleotide sequence as set forth in Figure 9, an attL1 nucleotide sequence as set forth in Figure 9, an attL2 nucleotide sequence as set forth in Figure 9, an attR1 nucleotide sequence as set forth in Figure 9, an attR2 nucleotide sequence as set forth in Figure 9, a polynucleotide complementary thereto, and a mutant, fragment, or derivative thereof.

2. An isolated nucleic acid molecule comprising an attB1 nucleotide sequence as set forth in Figure 9, a polynucleotide complementary thereto, or a mutant, fragment, variant or derivative thereof.

3. An isolated nucleic acid molecule comprising an attB2 nucleotide sequence as set forth in Figure 9, a polynucleotide complementary thereto, or a mutant, fragment, variant or derivative thereof.

4. An isolated nucleic acid molecule comprising an attP1 nucleotide sequence as set forth in Figure 9, a polynucleotide complementary thereto, or a mutant, fragment, variant or derivative thereof.

5. An isolated nucleic acid molecule comprising an attP2 nucleotide sequence as set forth in Figure 9, a polynucleotide complementary thereto, or a mutant, fragment, variant or derivative thereof.

6. An isolated nucleic acid molecule comprising an attL1 nucleotide sequence as set forth in Figure 9, a polynucleotide complementary thereto, or a mutant, fragment, variant or derivative thereof.

7. An isolated nucleic acid molecule comprising an attL2 nucleotide sequence as set forth in Figure 9, a polynucleotide complementary thereto, or a mutant, fragment, variant or derivative thereof.

5 8. An isolated nucleic acid molecule comprising an attR1 nucleotide sequence as set forth in Figure 9, a polynucleotide complementary thereto, or a mutant, fragment, variant or derivative thereof.

10 9. An isolated nucleic acid molecule comprising an attR2 nucleotide sequence as set forth in Figure 9, a polynucleotide complementary thereto, or a mutant, fragment, variant or derivative thereof.

15 10. The isolated nucleic acid molecule of claim 1, further comprising one or more functional or structural nucleotide sequences selected from the group consisting of one or more multiple cloning sites, one or more localization signals, one or more transcription termination sites, one or more transcriptional regulatory sequences, one or more translational signals, one or more origins of replication, one or more fusion partner peptide-encoding nucleic acid molecules, one or more protease cleavage sites, and one or more 5' polynucleotide extensions.

20 11. The nucleic acid molecule of claim 10, wherein said transcriptional regulatory sequence is a promoter, an enhancer, or a repressor.

25 12. The nucleic acid molecule of claim 10, wherein said fusion partner peptide-encoding nucleic acid molecule encodes glutathione S-transferase (GST), hexahistidine (His₆), or thioredoxin (Trx).

30 13. The nucleic acid molecule of claim 10, wherein said 5' polynucleotide extension consists of from one to five nucleotide bases.

14. The nucleic acid molecule of claim 13, wherein said 5' polynucleotide extension consists of four or five guanine nucleotide bases.

15. A primer nucleic acid molecule suitable for amplifying a target nucleotide sequence, comprising the isolated nucleic acid molecule of claim 1 or a portion thereof linked to a target-specific nucleotide sequence useful in amplifying said target nucleotide sequence.

16. The primer nucleic acid molecule of claim 15, wherein said primer comprises an attB1 nucleotide sequence having the sequence shown in Figure 9 or a portion thereof, or a polynucleotide complementary to the sequence shown in Figure 9 or a portion thereof.

17. The primer nucleic acid molecule of claim 15, wherein said primer comprises an attB2 nucleotide sequence having the sequence shown in Figure 9 or a portion thereof, or a polynucleotide complementary to the sequence shown in Figure 9 or a portion thereof.

18. The primer nucleic acid molecule of claim 15, further comprising a 5' terminal extension of four or five guanine bases.

19. A vector comprising the isolated nucleic acid molecule of claim 1.

20. The vector of claim 19, wherein said vector is an Expression Vector.

21. A host cell comprising the isolated nucleic acid molecule of claim 1 or the vector of claim 19.

22. A method of synthesizing or amplifying one or more nucleic acid molecules comprising:

- (a) mixing one or more nucleic acid templates with at least one polypeptide having polymerase or reverse transcriptase activity and at least a first primer comprising a template-specific sequence that is complementary to or capable of hybridizing to said

-171-

templates and at least a second primer comprising all or a portion of a recombination site wherein said at least a portion of said second primer is homologous to or complementary to at least a portion of said first primer; and

- 5 (b) incubating said mixture under conditions sufficient to synthesize or amplify one or more nucleic acid molecules complementary to all or a portion of said templates and comprising one or more recombination sites or portions thereof at one or both termini of said molecules.

10 23. A method of synthesizing or amplifying one or more nucleic acid molecules comprising:

- 15 (a) mixing one or more nucleic acid templates with at least one polypeptide having polymerase or reverse transcriptase activity and at least a first primer comprising a template-specific sequence that is complementary to or capable of hybridizing to said templates and at least a portion of a recombination site, and at least a second primer comprising all or a portion of a recombination site wherein said at least a portion of said recombination site on said second primer is complementary to or homologous to at least a portion of said recombination site on said first primer; and

- 20 (b) incubating said mixture under conditions sufficient to synthesize or amplify one or more nucleic acid molecules complementary to all or a portion of said templates and comprising one or more recombination sites or portions thereof at one or both termini of said molecules.

25 24. A method of amplifying or synthesizing one or more nucleic acid molecules comprising:

- 30 (a) mixing one or more nucleic acid templates with at least one polypeptide having polymerase or reverse transcriptase activity

and one or more first primers comprising at least a portion of a recombination site and a template-specific sequence that is complementary to or capable of hybridizing to said template;

5 (b) incubating said mixture under conditions sufficient to synthesize or amplify one or more first nucleic acid molecules complementary to all or a portion of said templates wherein said molecules comprise at least a portion of a recombination site at one or both termini of said molecules;

10 (c) mixing said molecules with one or more second primers comprising one or more recombination sites, wherein said recombination sites of said second primers are homologous to or complementary to at least a portion of said recombination sites on said first nucleic acid molecules; and

15 (d) incubating said mixture under conditions sufficient to synthesize or amplify one or more second nucleic acid molecules complementary to all or a portion of said first nucleic acid molecules and which comprise one or more recombination sites at one or both termini of said molecules.

20 25. A polypeptide encoded by the isolated nucleic acid molecule of any one of claims 1-10.

25 26. An isolated nucleic acid molecule comprising one or more *att* recombination sites comprising at least one mutation in its core region that increases the specificity of interaction between said recombination site and a second *att* recombination site.

30 27. The isolated nucleic acid molecule of claim 26, wherein said mutation is at least one substitution mutation of at least one nucleotide in the seven basepair overlap region of said core region of said recombination site.

28. The isolated nucleic acid molecule of claim 26, wherein said nucleic acid molecule comprises the sequence NNNATAC, wherein "N" refers to any nucleotide with the proviso that if one of the first three nucleotides in the consensus sequence is a T/U, then at least one of the other two of the first three nucleotides is not a T/U.

29. An isolated nucleic acid molecule comprising one or more mutated *att* recombination sites comprising at least one mutation in its core region that enhances the efficiency of recombination between a first nucleic acid molecule comprising said mutated *att* recombination site and a second nucleic acid molecule comprising a second recombination site that interacts with said mutated *att* recombination site.

30. The isolated nucleic acid molecule of claim 29, wherein said mutated *att* recombination site is a mutated *attL* site comprising a core region having the nucleotide sequence caactnntnnnannaagttg, wherein "n" represents any nucleotide.

31. The isolated nucleic acid molecule of claim 30, wherein said mutated *attL* recombination site comprises a core region having a nucleotide sequence selected from agcctgctttattataactaagttggcatta (*attL5*) and agcctgctttttatattaagttggcatta (*attL6*).

32. The isolated nucleic acid molecule of claim 29, wherein said mutated *att* recombination site comprises a core region having a nucleotide sequence selected from the group consisting of ggggacaactttgtacaaaaagttggct (*attB1.6*), ggggacaactttgtacaagaagctgggt (*attB2.2*), and ggggacaactttgtacaagaagttgggt (*attB2.10*).

33. A vector selected from the group consisting of pENTR1A, pENTR2B, pENTR3C, pENTR4, pENTR5, pENTR6, pENTR7, pENTR8, pENTR9, pENTR10, pENTR11, pDEST1, pDEST2, pDEST3, pDEST4,

pDEST5, pDEST6, pDEST7, pDEST8, pDEST9, pDEST10, pDEST11, pDEST12.2 (also known as pDEST12), pDEST13, pDEST14, pDEST15, pDEST16, pDEST17, pDEST18, pDEST19, pDEST20, pDEST21, pDEST22, pDEST23, pDEST24, pDEST25, pDEST26, pDEST27, pDEST28, pDEST29, pDEST30, pDEST31, pDEST32, pDEST33, pDEST34, pDONR201 (also known as pENTR21 attP vector or pAttPkan Donor Vector), pDONR202, pDONR203 (also known as pEZ15812), pDONR204, pDONR205, pDONR206 (also known as pENTR22 attP vector or pAttPgen Donor Vector), pDONR207, pMAB58, pMAB62, pMAB85 and pMAB86.

34. A host cell comprising the vector of claim 33.

35. A polypeptide encoded by the vector of claim 33.

36. A kit for use in synthesizing a nucleic acid molecule, said kit comprising the isolated nucleic acid molecule of any one of claims 1-10, 26 and 29.

37. A kit for use in synthesizing a nucleic acid molecule, said kit comprising the primer of claim 15 or claim 18.

38. A kit for use in cloning a nucleic acid molecule, said kit comprising the vector of claim 19 or claim 33.

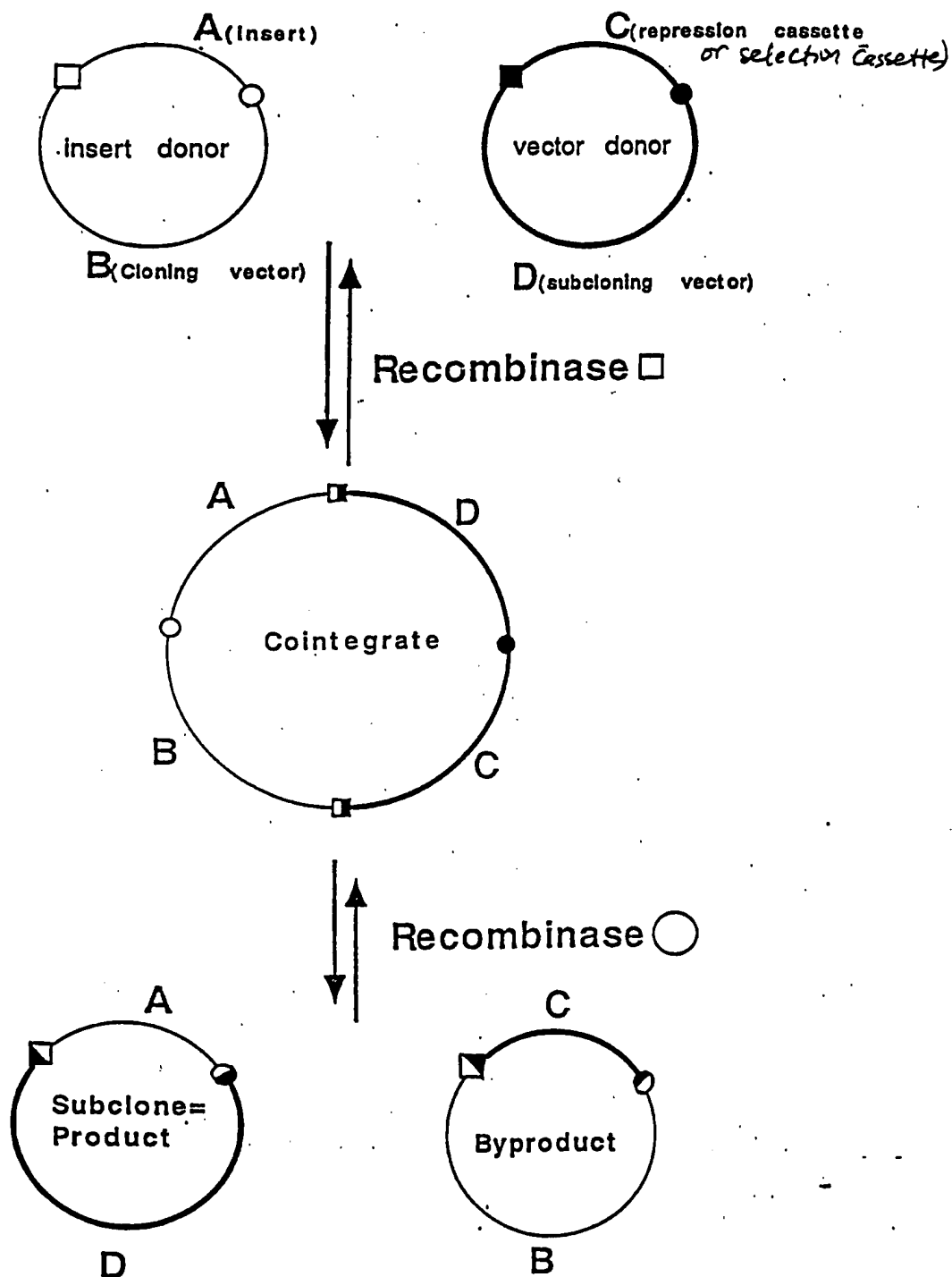


Figure 1

2/240

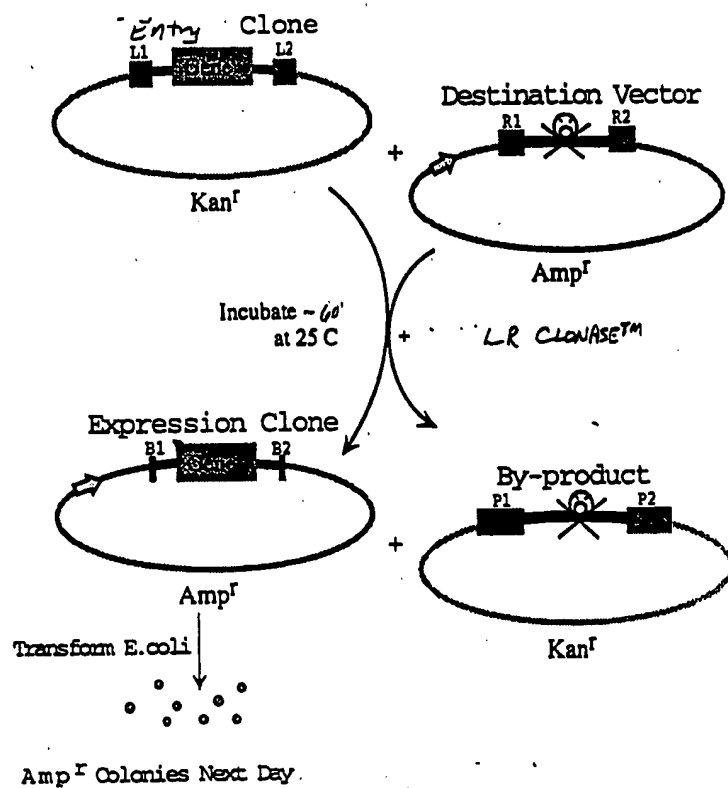


FIGURE 2

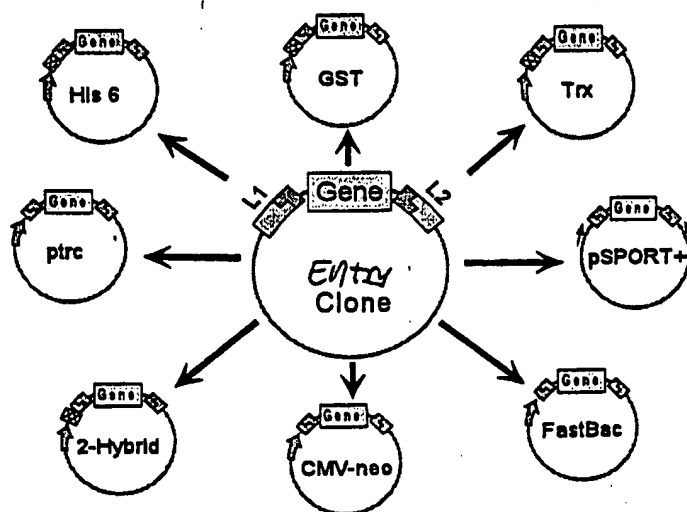


FIGURE 3

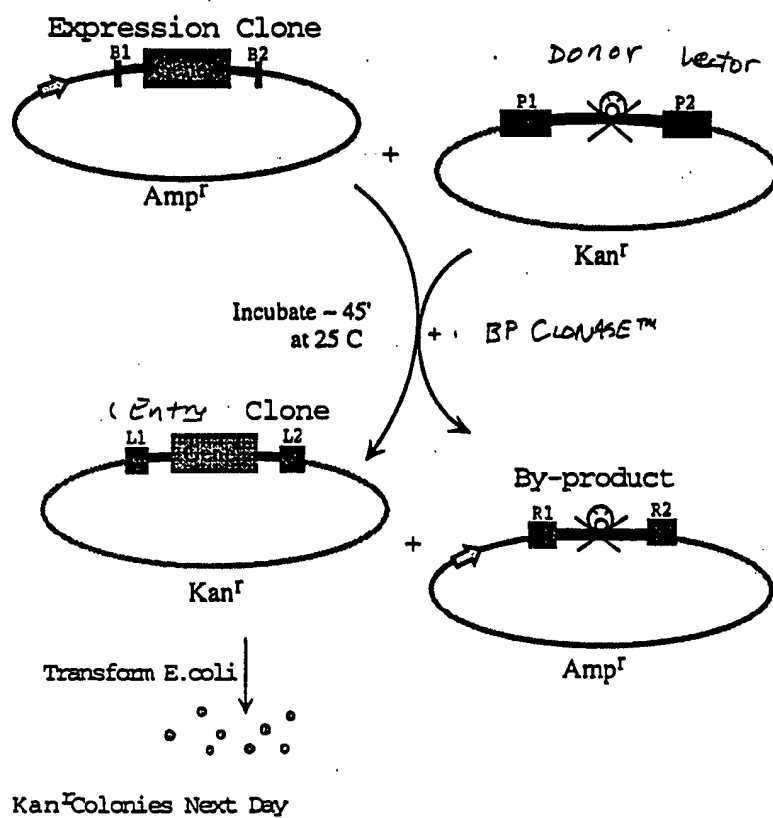


FIGURE 4

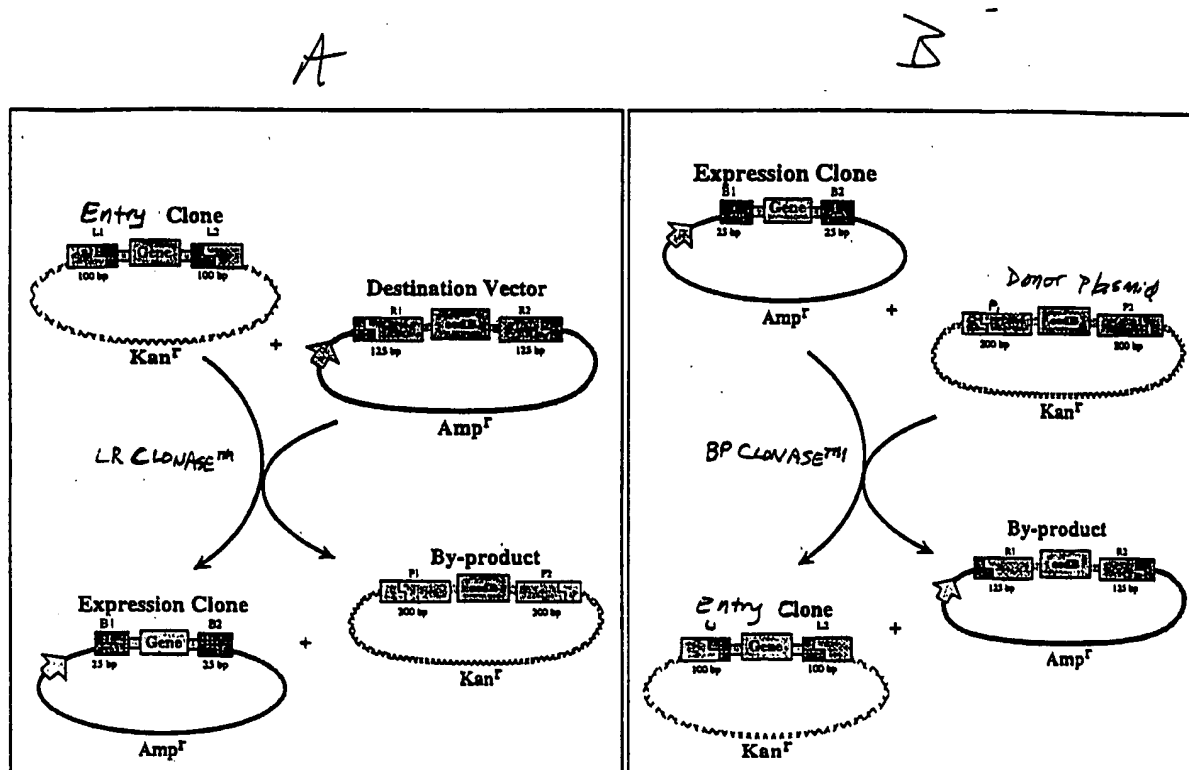


FIGURE 5

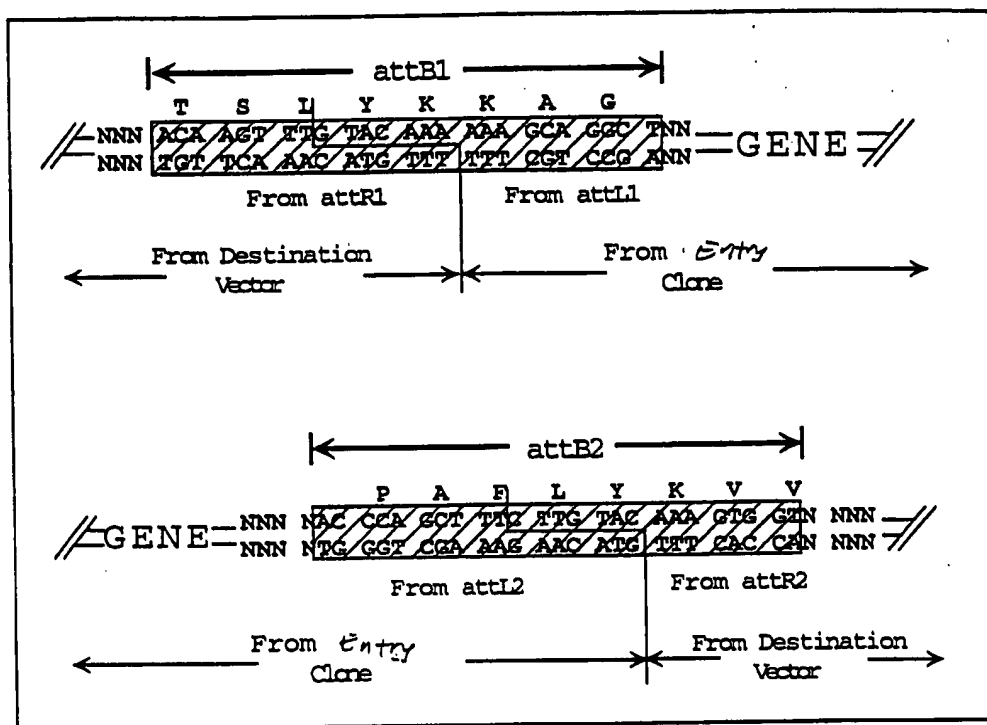


FIGURE 6

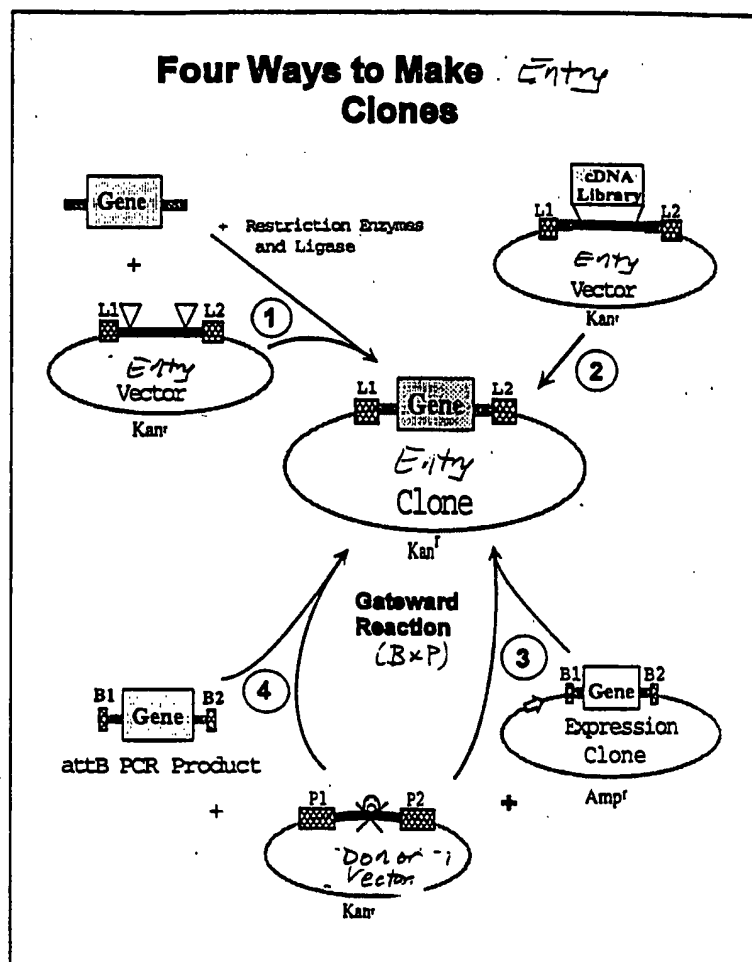


FIGURE 7

8/240

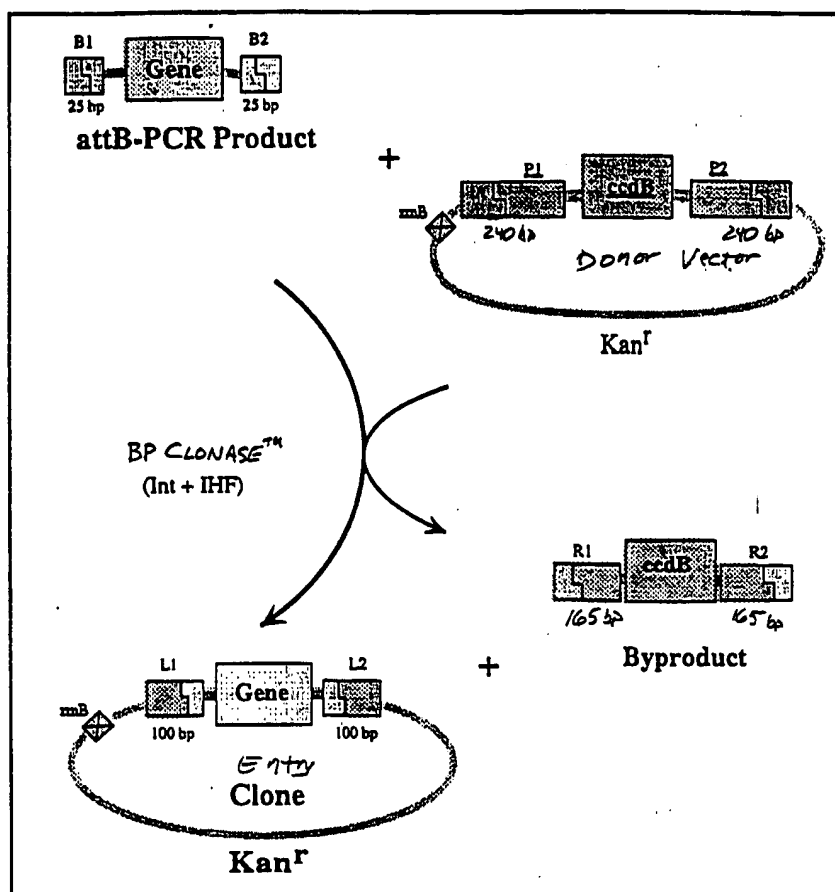


FIGURE 8

Recombination Site Nucleotide Sequences

attB1: 5'-ACAAGTTTGTACAAAAAGCAGGCT-3'

attB2: 5'-ACCCAGCTTTCTTGTACAAAGTGGT-3'

attP1: 5'-TACAGGTCACCTAATACCATCTAAGTAGTTGATTCATAGTGACTGGATATG-
TTGTGTTTTACAGTATTATGTAGTCTGTTTTTATGCAAAATCTAATTTA-
ATATATTGATATTTATATCATTTTACGTTTCTCGTTCAGCTTTTTGTAC-
AAAGTTGGCATTATAAAAAAGCATTGCTCATCAATTTGTTGCAACGAACA-
GGTCACTATCAGTCAAAATAAAATCATTATTTG-3'

attP2: 5'-CAAATAATGATTTTATTTTGACTGATAGTGACCTGTTGTTGCAACAAAT-
TGATAAGCAATGCTTTCTTATAATGCCAACTTTGTACAAGAAAGCTGAAC-
GAGAAACGTAAAATGATATAAATATCAATATATTAAATTAGATTTTGCAT-
AAAAACAGACTACATAATACTGTAAACACAACATATCCAGTCACTATGA-
ATCAACTACTTAGATGGTATTAGTGACCTGTA-3'

attR1: 5'-ACAAGTTTGTACAAAAAGCTGAACGAGAAACGTAAAATGATATAAA-
TATCAATATATTAAATTAGATTTTGCATAAAAAACAGACTACATAATAC-
TGTAACACACAACATATCCAGTCACTATG-3'

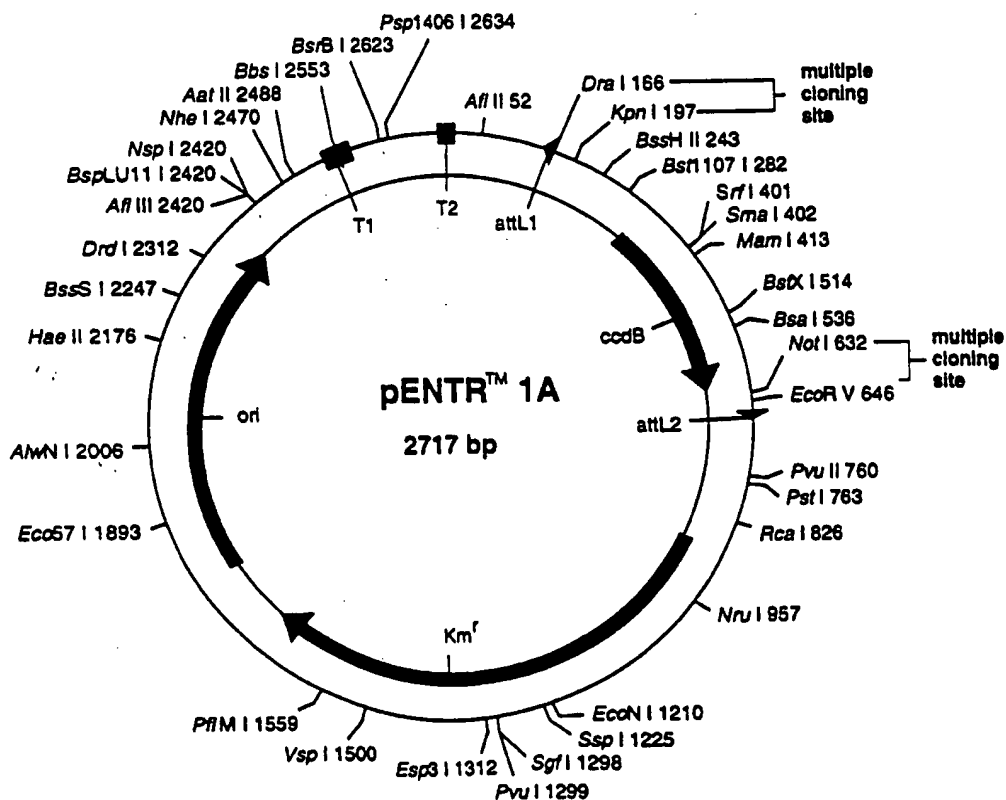
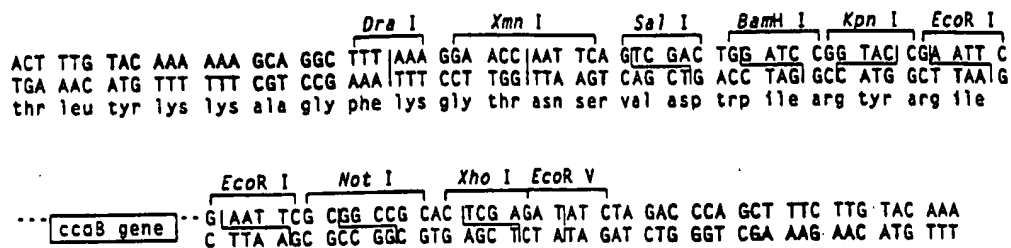
attR2: 5'-GCAGGTCGACCATAGTGACTGGATATGTTGTGTTTTACAGTATTAT-
GTAGTCTGTTTTTATGCAAAATCTAATTTAATATATTGATATTT-
ATATCATTTTACGTTTCTCGTTCAGCTTTCTTGTACAAAGTGGT-3'

attL1: 5'-CAAATAATGATTTTATTTTGACTGATAGTGACCTGTTGTTGCAAC-
AAATTGATAAGCAATGCTTTTTTATAATGCCAACTTTGTACAAAAA-
GCAGGCT-3'

attL2: 5'-CAAATAATGATTTTATTTTGACTGATAGTGACCTGTTGTTGCAACAA-
ATTGATAAGCAATGCTTTCTTATAATGCCAACTTTGTACAAGAAAGCTGGGT-3'

Figure 9

Figure 10A: Cloning sites of the Entry Vector pENTRTM 1A (reading frame A)



11/240

pENTR1A 2717 bp

<u>Base Nos.</u>	<u>Gene Encoded</u>
67..166	attL1
321..626	ccdB
655..754	attL2
877..1686	KmR
1791..2364	ori

```

1 CTGACGGATG GCCTTTTTCG GTTTCTACAA ACTCTTCCTG TTAGTTAGTT ACTTAAGCTC
61 GGGCCCCAAA TAATGATTTT ATTTTGACTG ATAGTGACCT GTTCGTTGCA ACAAATTGAT
121 AAGCAATGCT TTTTATAAT GCCAACTTTG TACAAAAAAG CAGGCTTTAA AGGAACCAAT
181 TCAGTCGACT GGATCCGGTA CCGAATTCGC TTACTAAAAG CCAGATAACA GTATGCGTAT
241 TTGCGCGCTG ATTTTTCGCG TATAAGAATA TATACTGATA TGTATACCCG AAGTATGTCA
301 AAAAGAGGTG TGCTTCTAGA ATGCAGTTTA AGGTTTACAC CTATAAAAGA GAGAGCCGTT
361 ATCGTCTGTT TGTGGATGTA CAGAGTGATA TTATTGACAC GCCCGGGCGA CGGATAGTGA
421 TCCCCCTGGC CAGTGCACGT CTGCTGTCAG ATAAAGTCTC CCGTGAACCT TACCCGGTGG
481 TGCAATATCGG GGATGAAAGC TGGCGCATGA TGACCACCGA TATGGCCAGT GTGCCGGTCT
541 CCGTTATCGG GGAAGAAGTG GCTGATCTCA GCCACCGCGA AAATGACATC AAAAAACGCCA
601 TTAACCTGAT GTTCTGGGGA ATATAGAATT CGCGGCCGCA CTCGAGATAT CTAGACCCAG
661 CTTTCTTGTA CAAAGTTGGC ATTATAAGAA AGCATTGCTT ATCAATTTGT TGCAACGAAC
721 AGGTCACTAT CAGTCAAAAT AAAATCATT TTTGCCATCC AGCTGCAGCT CTGGCCCGTG
781 TCTCAAAATC TCTGATGTTA CATTGCACAA GATAAAAATA TATCATCATG AACAATAAAA
841 CTGTCTGCTT ACATAAACAG TAATACAAAG GGTGTTATGA GCCATATTCA ACGGGAAACG
901 TCGAGGCCGC GATTAAATTC CAACATGGAT GCTGATTAT ATGGGTATAA ATGGGCTCGC
961 GATAATGTCG GGCAATCAGG TGGGACAATC TATCGCTTGT ATGGGAAGCC CGATGCGCCA
1021 GAGTGTGTTT TGAAACATGG CAAAGGTAGC GTTGCCAATG ATGTTACAGA TGAGATGGTC
1081 AGACTAAACT GGCTGACGGA ATTTATGCCT CTTCCGACCA TCAAGCATT TATCCGTACT
1141 CCTGATGATG CATGGTTACT CACCACTGCG ATCCCGGAA AAACAGCATT CCAGGTATTA
1201 GAAGAATATC CTGATTCAGG TGAAAATATT GTTGATGCGC TGGCAGTGTC CCTGCGCCGG
1261 TTGCATTGTA TTCCTGTTTG TAATTGTCCT TTTAACAGCG ATCGCGTATT TCGTCTCGCT
1321 CAGGCGCAAT CACGAATGAA TAACGGTTTG GTTGATGCGA GTGATTTTGA TGACGAGCGT
1381 AATGGCTGGC CTGTTGAACA AGTCTGGAAG GAAATGCATA AACTTTTGCC ATTCTCACCG
1441 GATTCACTCG TCACTCATGG TGATTTCTCA CTTGATAACC TTATTTTTGA CGAGGGGAAA
1501 TTAATAGGTT GTATTGATGT TGGACGAGTC GGAATCGCAG ACCGATACCA GGATCTTGCC
1561 ATCCTATGGA ACTGCCTCGG TGAGTTTCTT CCTTCATTAC AGAAACGGCT TTTTCAAAAA
1621 TATGGTATTG ATAATCCTGA TATGAATAAA TTGCAGTTTC ATTTGATGCT CGATGAGTTT
1681 TTCTAATCAG AATTGGTTAA TTGGTTGTAA CATTATTCAG ATTGGGCCCC GTTCCACTGA
1741 GCGTCAGACC CCGTAGAAAA GATCAAAGGA TCTTCTTGAG ATCCTTTTTT TCTGCGCGTA
1801 ATCTGCTGCT TGCAACAAAA AAAACCACCG CTACCAGCGG TGGTTTGTTC GCCGGATCAA
1861 GAGCTACCAA CTCTTTTTCG GAAGGTAAGT GGCTTCAGCA GAGCGCAGAT ACCAAATACT
1921 GTTCTTCTAG TGTAGCCGTA GTTAGGCCAC CACTTCAAGA ACTCTGTAGC ACCGCCTACA
1981 TACCTCGCTC TGCTAATCCT GTTACCAGTG GCTGCTGCCA GTGGCGATAA GTCGTGTCTT
2041 ACCGGGTTGG ACTCAAGACG ATAGTTACCG GATAAGGCGC AGCGGTCGGG CTGAACGGGG
2101 GGTTCGTGCA CACAGCCCAG CTTGGAGCGA ACGACCTACA CCGAACTGAG ATACCTACAG
2161 CGTGAGCTAT GAGAAAGCGC CACGCTTCCC GAAGGGAGAA AGGCGGACAG GTATCCGGTA
2221 AGCGGCAGGG TCGGAACAGG AGAGCGCACG AGGGAGCTTC CAGGGGGAAA CGCCTGGTAT
2281 CTTTATAGTC CTGTCGGGTT TCGCCACCTC TGACTTGAGC GTCGATTTTT GTGATGCTCG
2341 TCAGGGGGGC GGAGCCTATG GAAAAACGCC AGCAACGCGG CCTTTTTACG GTTCTGGGCC
2401 TTTTGCTGGC CTTTGTCTCA CATGTTCTTT CCTGCGTTAT CCCCTGATTC TGTGGATAAC
2461 CGTATTACCG CTAGCATGGA TCTCGGGGAC GTCTAACTAC TAAGCGAGAG TAGGGAAGTT
2521 CCAGGCATCA AATAAAACGA AAGGCTCAGT CGGAAGACTG GGCCTTTCGT TTTATCTGTT
2581 GTTGTGCGGT GAACGCTCTC CTGAGTAGGA CAAATCCGCC GGGAGCGGAT TTGAACGTTG
2641 TGAAGCAACG GCCCGGAGGG TGGCGGGCAG GACGCCCGCC ATAAACTGCC AGGCATCAAA
2701 CTAAGCAGAA GGCCATC

```

FIGURE 10B

Figure 11A: Cloning Sites of the Entry Vector pENTR2B (reading frame B)

Int	attL1	EheI	XmnI	SalI	BamHI
TTG	TAC AAA AAA GCA GGC TGG	CFC	CGG AAC CAA	TTC AGT CGA	CTG GAT CCG
AAC	ATG TTT TTT CGT CCG ACC	GCG	GCC TTG GTT	AAG TCA GCT	GAC CTA GGC
Leu	Tyr Lys Lys Ala Gly Trp	Arg Arg Asn Gln	Phe Ser Arg	Leu Asp Pro	

KpnI	EcoRI	EcoRI	NotI	XhoI	EcoRV	XbaI
GTA CCG AAT TC-	ccdB	--G AAT TCG CCG CCG	CAC TCG AGA TAT	CTA GAC CCA		
CAT GGC TTA AG		C TTA AGC GCC GGC	GTG AGC TCT ATA	GAT CTG GGT		
Val Pro Asn		Asn Ser Arg Pro	His Ser Arg Tyr	Leu Asp Pro		

Int	attL2
GCT TTC TTG TAC AAA G	
CGA AAG AAC ATG TTT C	
Ala Phe Leu Tyr Lys	

pENTR2B 2718 bp

<u>Location (Base Nos.)</u>	<u>Gene Encoded</u>
67..166	attL1
322..627	ccdB
656..755	attL2
878..1687	KmR
1792..2365	ori

```

1 CTGACGGATG GCCTTTTTCG GTTCTACAA ACTCTTCCTG TTAGTTAGTT ACTTAAGCTC
61 GGGCCCCAAA TAATGATTTT ATTTTGACTG ATAGTGACCT GTTCGTTGCA ACAAAATGAT
121 AAGCAATGCT TTTTATAAT GCCAACTTTG TACAAAAAAG CAGGCTGGCG CCGGAACCAA
181 TTCAGTCGAC TGGATCCGGT ACCGAATTCG CTTACTAAAA GCCAGATAAC AGTATGCGTA
241 TTTGCGCGCT GATTTTTCG GTATAAGAAT ATATACTGAT ATGTATACCC GAAGTATGTC
301 AAAAAGAGGT GTGCTTCTAG AATGCAGTTT AAGGTTTACA CCTATAAAAG AGAGAGCCGT
361 TATCGTCTGT TTGTGGATGT ACAGAGTGAT ATTATTGACA CGCCCGGGCG ACGGATGGTG
421 ATCCCCCTGG CCAGTGCACG TCTGCTGTCA GATAAAGTCT CCCGTGAACT TTACCCGGTG
481 GTGCATATCG GGGATGAAAG CTGGCGCATG ATGACCACCG ATATGGCCAG TGTGCCGGTC
541 TCCGTTATCG GGGAAGAAGT GGCTGATCTC AGCCACCGCG AAAATGACAT CAAAACGCC
601 ATTAACCTGA TGTTCGGGG AATATAGAAT TCGCGGCCGC ACTCGAGATA TCTAGACCCA
661 GCTTCTTGT ACAAGTTGG CATTATAAGA AAGCATTGCT TATCAATTG TTGCAACGAA
721 CAGGTCACCTA TCAGTCAAAA TAAATCATT ATTTGCCATC CAGCTGCAGC TCTGGCCCGT
781 GTCTCAAAAT CTCTGATGTT ACATTGCACA AGATAAAAAAT ATATCATCAT GAACAATAAA
841 ACTGTCGTCT TACATAAACA GTAATACAAG GGGTGTATG AGCCATATT AACGGGAAAC
901 GTCGAGGCCG CGATTAAATT CCAACATGGA TGCTGATTTA TATGGGTATA AATGGGCTCG
961 CGATAATGTC GGGCAATCAG GTGCGACAAT CTATCGCTTG TATGGGAAGC CCGATGCGCC
1021 AGAGTTGTTT CTGAAACATG GCAAAGGTAG CGTTGCCAAT GATGTTACAG ATGAGATGGT
1081 CAGACTAAAC TGGCTGACGG AATTTATGCC TCTTCCGACC ATCAAGCATT TTATCCGTAC
1141 TCCTGATGAT GCATGGTTAC TCACCACTGC GATCCCCGGA AAAACAGCAT TCCAGGTATT
1201 AGAAGAATAT CCTGATTGAG GTGAAAATAT TGTGATGCG CTGGCAGTGT TCCTGCGCCG
1261 GTTGCAATCG ATTCCTGTTT GTAATTGTCC TTTTAACAGC GATCGCGTAT TTCGTCTCGC
1321 TCAGGCGCAA TCACGAATGA ATAACGGTTT GGTGATGCG AGTGATTTTG ATGACGAGCG
1381 TAATGGCTGG CCTGTTGAAC AAGTCTGGAA AGAAATGCAT AAACCTTTGC CATTCTCACC
1441 GGATTCAGTC GTCACATG GTGATTTCTC ACTTGATAAC CTTATTTTTC ACGAGGGGAA
1501 ATTAATAGGT TGTATTGATG TTGACGAGT CGGAATCGCA GACCGATACC AGGATCTTGC
1561 CATCCTATGG AACTGCCTCG GTGAGTTTTC TCCTTCATTA CAGAAACGGC TTTTCAAAA
1621 ATATGGTATT GATAATCCTG ATATGAATAA ATTGCAGTTT CATTTGATGC TCGATGAGTT
1681 TTTCTAATCA GAATTGGTTA ATTGGTTGTA ACATTATTCA GATTGGGCCC CGTTCCACTG
1741 AGCGTCAGAC CCCGTAGAAA AGATCAAAGG ATCTTCTTGA GATCCTTTTT TTCTGCGCGT
1801 AATCTGCTGC TTGCAACAA AAAAACCACC GCTACCAGCG GTGGTTTGTT TGCCGGATCA
1861 AGAGCTACCA ACTCTTTTTC CGAAGGTAAC TGGCTTCAGC AGAGCGCAGA TACCAAATAC
1921 TGTTCTTCTA GTGTAGCCGT AGTTAGGCCA CCACTTCAAG AACTCTGTAG CACCGCCTAC
1981 ATACCTCGCT CTGCTAATCC TGTACCAGT GGCTGCTGCC AGTGGCGATA AGTCGTGTCT
2041 TACCGGGTTG GACTCAAGAC GATAGTTACC GGATAAGGCG CAGCGGTCGG GCTGAACGGG
2101 GGGTTCGTGC ACACAGCCCA GCTTGGAGCG AACGACCTAC ACCGAACTGA GATACCTACA
2161 GCGTGAGCTA TGAGAAAGCG CCACGCTTCC CGAAGGGAGA AAGGCGGACA GGTATCCGGT
2221 AAGCGGCAGG GTCGGAACAG GAGAGCGCAC GAGGGAGCTT CCAGGGGAA ACGCCTGTA
2281 TCTTTATAGT CCTGTCGGGT TTCGCCACCT CTGACTTGAG CGTCGATTTT TGTGATGCTC
2341 GTCAGGGGGG CGGAGCCTAT GGAAAAACGC CAGCAACGCG GCCTTTTAC GGTTCCTGGC
2401 CTTTGTCTGG CCTTTTGCTC ACATGTTCTT TCCTGCGTTA TCCCCTGATT CTGTGGATAA
2461 CCGTATTACC GCTAGCATGG ATCTCGGGGA CGTCTAACTA CTAAGCGAGA CTAGGGAAC
2521 GCCAGGCATC AAATAAACG AAAGGCTCAG TCGGAAGACT GGGCCTTTCG TTTTATCTGT
2581 TGTTTGTCGG TGAACGCTCT CCTGAGTAGG ACAAATCCGC CGGGAGCGGA TTTGAACGTT
2641 GTGAAGCAAC GGCCCGGAGG GTGGCGGGCA GGACGCCGCG CATAACTG CAGGCATCAA
2701 ACTAAGCAGA AGGCCATC

```

FIGURE 11B

Figure 2A: Cloning Sites of the Entry Vector pENTR3C (reading frame C) ...

Int	attL1		DraI		XmnI		SalI		BamHI								
TTG	TAC	AAA	AAA	GCA	GGC	TCT	TTA	AAG	GAA	CCA	ATT	CAG	TCG	ACT	GGA	TCC	GGT
AAC	ATG	TTT	TTT	CGT	CCG	AGA	AAT	TTC	CTT	GGT	TAA	GTC	AGC	TGA	CCT	AGG	OCA
							↓			↓				↓		↓	↓
Leu	Tyr	Lys	Lys	Ala	Gly	Ser	Leu	Lys	Glu	Pro	Ile	Gln	Ser	Thr	Gly	Ser	Gly

KpnI	EcoRI		PvuI		EcoRI		NotI		XhoI		EcoRV	XbaI			
ACC	GAA	TTC	GAT	GSC--	ccdB	--G	AAT	TCG	CGG	CCG	CAC	TCG	AGA	TAT	CTA
TGG	CTT	AAG	CTA	GCG		C	TTA	AGC	GCC	GGC	GTG	AGC	TCT	ATA	GAT
			↓				↓		↓		↓		↓		
Thr	Glu	Phe					Asn	Ser	Arg	Pro	His	Ser	Arg	Tyr	Leu

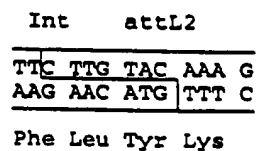
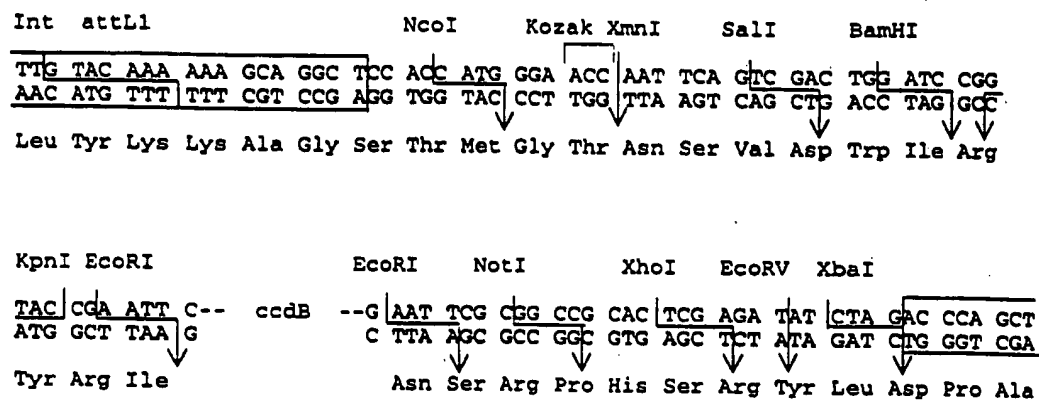
attL2		Int					
GAC	CCA	GCT	TTT	TTG	TAC	AAA	G
CTG	GGT	CGA	AAG	AAC	ATG	TTT	C
			↓				
Asp	Pro	Ala	Phe	Leu	Tyr	Lys	

15/240

pENTR3C 2723 bp

<u>Location (Base Nos.)</u>		<u>Gene Encoded</u>
67..166		attL1
327..632		ccdB
661..760		attL2
883..1692		KmR
1797..2370		ori
1	CTGACGGATG GCCTTTTTCG GTTTCTACAA ACTCTTCCTG TTAGTTAGTT ACTTAAGCTC	
61	GGGCCCCAAA TAATGATTTT ATTTTGACTG ATAGTGACCT GTTCGTTGCA ACAAATTGAT	
121	AAGCAATGCT TTTTATAAT GCCAATTTG TACAAAAAAG CAGGCTCTTT AAAGGAACCA	
181	ATTCAGTCGA CTGGATCCGG TACCGAATTC GATCGCTTAC TAAAAGCCAG ATAACAGTAT	
241	GCGTATTTGC GCGCTGATTT TTGCGGTATA AGAATATATA CTGATATGTA TACCCGAAGT	
301	ATGTCAAAAA GAGGTGTGCT TCTAGAATGC AGTTTAAGGT TTACACCTAT AAAAGAGAGA	
361	GCCGTTATCG TCTGTTTGTG GATGTACAGA GTGATATTAT TGACACGCCC GGGCGACGGA	
421	TGGTGATCCC CCTGGCCAGT GCACGCTGCG TGTCAGATAA AGTCTCCCGT GAACTTTACC	
481	CGGTGGTGCA TATCGGGGAT GAAAGCTGGC GCATGATGAC CACCGATATG GCCAGTGTGC	
541	CGGTCTCCGT TATCGGGGAA GAAGTGGCTG ATCTCAGCCA CCGCGAAAAT GACATCAAAA	
601	ACGCCATTAA CCTGATGTTT TGGGGAATAT AGAATTCGCG GCCGCACTCG AGATATCTAG	
661	ACCCAGCTTT CTGTACAAA GTTGGCATT TAAGAAAGCA TTGCTTATCA ATTTGTTGCA	
721	ACGAACAGGT CACTATCAGT CAAAATAAAA TCATTATTTG CCATCCAGT GCAGCTCTGG	
781	CCCGTGTCTC AAAATCTCTG ATGTTACATT GCACAAGATA AAAATATATC ATCATGAACA	
841	ATAAACTGT CTGCTTACAT AAACAGTAAT ACAAGGGGTG TTATGAGCCA TATTCAACGG	
901	GAAACGTCGA GGCCGCGATT AAATCCAAAC ATGGATGCTG ATTTATATGG GTATAAATGG	
961	GCTCGCGATA ATGTCGGGCA ATCAGGTGCG ACAATCTATC GCTTGATATGG GAAGCCCGAT	
1021	GCGCCAGAGT TGTTTCTGAA ACATGGCAAA GGTAGCGTTG CCAATGATGT TACAGATGAG	
1081	ATGGTCAGAC TAAACTGGCT GACGGAATTT ATGCCTCTTC CGACCATCAA GCATTTTATC	
1141	CGTACTCCTG ATGATGCATG GTTACTCACC ACTGCGATCC CCGGAAAAAC AGCATTCCAG	
1201	GTATTAGAAG AATATCCTGA TTCAGGTGAA AATATTGTTG ATGCGCTGGC AGTGTTCCTG	
1261	CGCCGGTTGC ATTCGATTCC TGTTTGTAAT TGTCCTTTTA ACAGCGATCG CGTATTTCTG	
1321	CTCGCTCAGG CGCAATCACG AATGAATAAC GGTTTGGTTG ATGCGAGTGA TTTTGATGAC	
1381	GAGCGTAATG GCTGGCCTGT TGAACAAGTC TGGAAAGAAA TGCATAAACT TTTGCCATT	
1441	TCACCGGATT CAGTCGTAC TCATGGTGAT TTCTCACTTG ATAACCTTAT TTTTGACGAG	
1501	GGGAAATTAA TAGGTTGTAT TGATGTTGGA CGAGTCGGAA TCGCAGACCG ATACCAGGAT	
1561	CTTGCCATCC TATGGAATG CCTCGGTGAG TTTTCTCCTT CATTACAGAA ACGGCTTTTT	
1621	CAAAAATATG GTATTGATAA TCCTGATATG AATAAATTGC AGTTTCATTG GATGCTCGAT	
1681	GAGTTTTTCT AATCAGAATT GGTAAATTGG TTGTAACATT ATTCAGATTG GGCCCCGTTT	
1741	CACTGAGCGT CAGACCCCGT AGAAAAGATC AAAGGATCTT CTTGAGATCC TTTTTTCTG	
1801	CGCGTAATCT GCTGCTTGCA AACAAAAAAA CCACCGCTAC CAGCGGTGGT TTGTTGCCG	
1861	GATCAAGAGC TACCAACTCT TTTTCCGAAG GTAACCTGGT TCAGCAGAGC GCAGATACCA	
1921	AATACTGTTT TTCTAGTGTA GCCGTAGTTA GGCCACCACT TCAAGAATC TGTAGCACCG	
1981	CCTACATACC TCGCTCTGCT AATCCTGTGA CCAGTGGCTG CTGCCAGTGG CGATAAGTCG	
2041	TGTCTTACCG GGTGGGACTC AAGACGATAG TTACCGGATA AGGCGCAGCG GTCGGGCTGA	
2101	ACGGGGGGTT CGTGACACA GCCCAGCTTG GAGCGAACGA CCTACACCGA ACTGAGATAC	
2161	CTACAGCGTG AGCTATGAGA AAGCGCCACG CTTCCCGAAG GGAGAAAGGC GGACAGGTAT	
2221	CCGGTAAGCG GCAGGGTCGG AACAGGAGAG CGCAGGAGGG AGCTTCCAGG GGGAAACGCC	
2281	TGGTATCTTT ATAGTCCTGT CGGGTTTCGC CACCTCTGAC TTGAGCGTCG ATTTTGTGA	
2341	TGCTCGTCAG GGGGGCGGAG CCTATGGAAA AACGCCAGCA ACGCGGCCTT TTTACGGTTC	
2401	CTGGCCTTTT GCTGGCCTTT TGCTCACATG TTCTTCTCTG CGTTATCCCC TGATTCTGTG	
2461	GATAACCGTA TTACCGCTAG CATGGATCTC GGGGACGTCT AACTACTAAG CGAGAGTAGG	
2521	GAACTGCCAG GCATCAAATA AAACGAAAGG CTCAGTCGGA AGACTGGGCC TTTCTGTTTA	
2581	TCGTGTGTTT GTCGGTGAAC GCTCTCTGTA GTAGGACAAA TCCGCCGGGA GCGGATTGTA	
2641	ACGTTGTGAA GCAACGGCCC GGAGGGTGGC GGGCAGGACG CCCGCCATAA ACTGCCAGGC	
2701	ATCAAATAA GCAGAAGGCC ATC	

FIGURE 12B

Figure 13A: Cloning Sites of the Entry Vector pENTR4 :

17/240

pENTR4 2720 bp

<u>Location (Base Nos.)</u>	<u>Gene Encoded</u>
67..166	attL1
324..629	ccdB
658..757	attL2
880..1689	KmR
1794..2367	ori

```

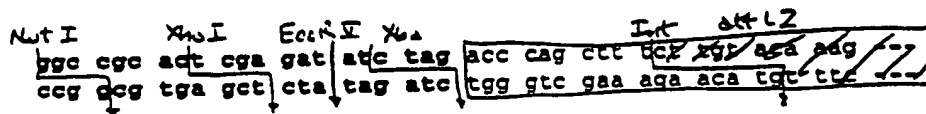
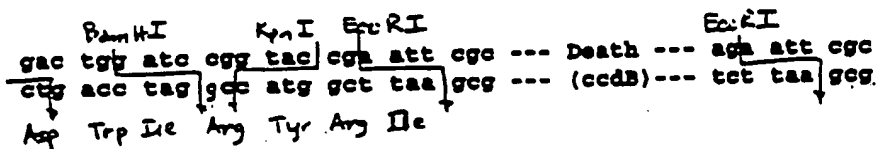
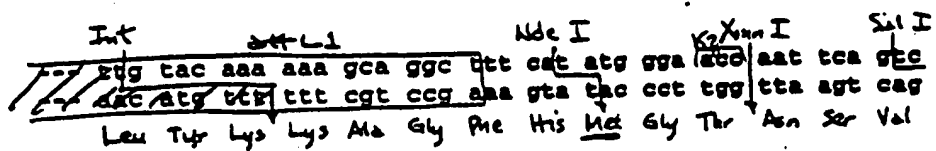
1 CTGACGGATG GCCTTTTTGC GTTTCTACAA ACTCTTCCTG TTAGTTAGTT ACTTAAGCTC
61 GGGCCCCAAA TAATGATTTT ATTTTGACTG ATAGTGACCT GTTCGTTGCA ACAAATTGAT
121 AAGCAATGCT TTTTATAAT GCCAAGTTG TACAAAAAG CAGGCTCCAC CATGGGAACC
181 AATTCAGTCG ACTGGATCCG GTACCGAATT CGCTTACTAA AAGCCAGATA ACAGTATGCG
241 TATTTGCGCG CTGATTTTTG CGGTATAAGA ATATATACTG ATATGTATAC CCGAAGTATG
301 TCAAAAAGAG GTGTGCTTCT AGAATGCAGT TTAAGGTTTA CACCTATAAA AGAGAGAGCC
361 GTTATCGTCT GTTTGTGGAT GTACAGAGTG ATATTATTGA CACGCCCGGG CGACGGATGG
421 TGATCCCCCT GGCCAGTGCA CGTCTGCTGT CAGATAAAGT CTCCCGTGAA CTTTACCCCG
481 TGGTGCAATAT CGGGGATGAA AGCTGGCGCA TGATGACCAC CGATATGGCC AGTGTGCCGG
541 TCTCCGTTAT CGGGGAAGAA GTGGCTGATC TCAGCCACCG CGAAAATGAC ATCAAAAACG
601 CCATTAACTT GATGTTCTGG GGAATATAGA ATTGCGGGCC GCACTCGAGA TATCTAGACC
661 CAGCTTTCTT GTACAAAGTT GGCATTATAA GAAAGCATTG CTTATCAATT TGTGCAACG
721 AACAGGTAC TATCAGTCAA AATAAATCA TTATTTGCCA TCCAGCTGCA GCTCTGGCCC
781 GTGTCTCAA ATCTCTGATG TTACATTGCA CAAGATAAAA ATATATCATC ATGAACAATA
841 AAAGTGTCTG CTTACATAAA CAGTAATACA AGGGGTGTTA TGAGCCATAT TCAACGGGAA
901 ACGTCGAGGC CGCGATTAAA TTCCAACATG GATGCTGATT TATATGGGTA TAAATGGGCT
961 CGCGATAATG TCGGGCAATC AGGTGCGACA ATCTATCGCT TGTATGGGAA GCCCGATGCG
1021 CCAGAGTTGT TTCTGAAACA TGGCAAAGGT AGCGTTGCCA ATGATGTTAC AGATGAGATG
1081 GTCAGACTAA ACTGGCTGAC GGAATTTATG CCTCTTCCGA CCATCAAGCA TTTTATCCGT
1141 ACTCCTGGTG ATGCATGGTT ACTCACCCT GCGATCCCCG GAAAAACAGC ATTCAGGTA
1201 TTAGAAGAAT ATCCTGATTC AGGTGAAAAT ATTGTTGATG CGCTGGCAGT GTTCCTGCGC
1261 CGGTTGCATT CGATTCTGTG TTGTAATTGT CCTTTAACA GCGATCGCGT ATTTCTGCTC
1321 GCTCAGGCGC AATCACGAAT GAATAACGGT TTGGTTGATG CGAGTGATTG TGATGACGAG
1381 CGTAATGGCT GGCCTGTTGA ACAAGTCTGG AAAGAAATGC ATAACTTTT GCCATTCTCA
1441 CCGGATTCAG TCGTCACTCA TGGTGATTTC TCACCTTGATA ACCTTATTTT TGACGAGGGG
1501 AAATTAATAG GTTGTATTGA TGTGAGACGA GTCGGAATCG CAGACCGATA CCAGGATCTT
1561 GCCATCCTAT GGAAGTGCCT CGGTGAGTTT TCTCCTTCAT TACAGAAACG GCTTTTTCAA
1621 AAATATGGTA TTGATAATCC TGATATGAAT AAATTGCAGT TTCATTTGAT GCTCGATGAG
1681 TTTTCTAAT CAGAATTGGT TAATTGGTTG TAACATTATT CAGATTGGGC CCCGTTCCAC
1741 TGAGCGTCAG ACCCGTAGA AAAGATCAA GGATCTTCTT GAGATCCTTT TTTTCTGCGC
1801 GTAATCTGCT GCTTGCAAAC AAAAAACCA CCGCTACCAG CGGTGGTTTG TTTGCCGGAT
1861 CAAGAGCTAC CAACTCTTTT TCCGAAGGTA ACTGGCTTCA GCAGAGCGCA GATACCAAAT
1921 ACTGTTCTTC TAGTGTAGCC GTAGTTAGGC CACCCTTCA AGAACTCTGT AGCACCCTCT
1981 ACATACCTCG CTCTGCTAAT CCTGTTACCA GTGGCTGCTG CCAGTGGCGA TAAGTCGTGT
2041 CTTACCGGGT TGGACTCAAG ACGATAGTTA CCGGATAAGG CGCAGCGGTC GGGCTGAACG
2101 GGGGGTTCGT GCACACAGCC CAGCTTGGAG CGAACGACCT ACACCGAACT GAGATACCTA
2161 CAGCGTGAGC TATGAGAAAG CGCCACGCTT CCCGAAGGGA GAAAGGCGGA CAGGTATCCG
2221 GTAAGCGGCA GGGTCGGAAC AGGAGAGCGC ACGAGGGAGC TTCCAGGGGG AAACGCTTGG
2281 TATCTTTATA GTCTGTGCGG GTTTCGCCAC CTCTGACTTG AGCGTCGATT TTTGTGATGC
2341 TCGTCAGGGG GCGGAGCCT ATGGAAAAAC GCCAGCAACG CGGCCTTTTT ACGGTTCTCTG
2401 GCCTTTTGCT GGCCTTTTGC TCACATGTTT TTTCTGCGT TATCCCTGA TTCTGTGGAT
2461 AACCGTATTA CCGCTAGCAT GGATCTCGGG GACGTCTAAC TACTAAGCGA GAGTAGGGAA
2521 CTGCCAGGCA TCAAAATAAA CGAAAGGCTC AGTCGGAAGA CTGGGCCTTT CGTTTTATCT
2581 GTTGTGTTGTC GGTGAACGCT CTCCTGAGTA GGACAAATCC GCCGGGAGCG GATTGAAACG
2641 TTGTGAAGCA ACGGCCCCGA GGGTGGCGGG CAGGACGCCC GCCATAAACT GCCAGGCATC
2701 AAAGTAAGCA GAAGGCCATC

```

FIGURE 13B

18/240

Figure 14A: Cloning sites of the Entry Vector pENTR5



19/240

pENTR5 2720 bp

<u>Location (Base Nos.)</u>	<u>Gene Encoded</u>
67..166	attL1
324..629	ccdB
658..757	attL2
880..1689	KmR
1794..2367	ori

```

1 CTGACGGATG GCCTTTTTCG GTTCTACAA ACTCTTCCTG TTAGTTAGTT ACTTAAGCTC
61 GGGCCCCAAA TAATGATTTT ATTTTGACTG ATAGTGACCT GTTCGTTGCA ACAAATTGAT
121 AAGCAATGCT TTTTATAAT GCCAACTTTG TACAAAAAAG CAGGCTTTCA TATGGGAACC
181 AATTCAGTCG ACTGGATCCG GTACCGAATT CGCTTACTAA AAGCCAGATA ACAGTATGCG
241 TATTTGCGCG CTGATTTTTCG CGGTATAAGA ATATATACTG ATATGTATAC CCGAAGTATG
301 TCAAAAAGAG GTGTGCTTCT AGAATGCAGT TTAAGGTTTA CACCTATAAA AGAGAGAGCC
361 GTTATCGTCT GTTTGTGGAT GTACAGAGTG ATATTATTGA CACGCCCGGG CGACGGATGG
421 TGATCCCCCT GGCCAGTGCA CGTCTGCTGT CAGATAAAGT CTCCCGTGAA CTTTACCCGG
481 TGGTGCATAT CGGGGATGAA AGCTGGCGCA TGATGACCAC CGATATGGCC AGTGTGCCGG
541 TCTCCGTTAT CGGGGAAGAA GTGGCTGATC TCAGCCACCG CGAAAATGAC ATCAAAAACG
601 CCATTAACTT GATGTTCTGG GGAATATAGA ATTCGCGGCC GCACTCGAGA TATCTAGACC
661 CAGCTTTCTT GTACAAAGTT GGCATTATAA GAAAGCATTG CTTATCAATT TGTGCAACG
721 AACAGGTCAC TATCAGTCAA AATAAAATCA TTATTGCCA TCCAGCTGCA GCTCTGGCCC
781 GTGTCTCAAA ATCTCTGATG TTACATTGCA CAAGATAAAA ATATATCATC ATGAACAATA
841 AAAGTGTCTG CTTACATAAA CAGTAATACA AGGGGTGTTA TGAGCCATAT TCAACGGGAA
901 ACGTCGAGGC CGCGATTAAA TTCCAACATG GATGCTGATT TATATGGGTA TAAATGGGCT
961 CGCGATAATG TCGGGCAATC AGGTGCGACA ATCTATCGCT TGTATGGGAA GCCCGATGCG
1021 CCAGAGTTGT TTCTGAAACA TGGCAAAGGT AGCGTTGCCA ATGATGTTAC AGATGAGATG
1081 GTCAGACTAA ACTGGCTGAC GGAATTTATG CCTCTTCCGA CCATCAAGCA TTTTATCCGT
1141 ACTCCTGATG ATGCATGGTT ACTCACCCT ACTCACCCT GCGATCCCCG GAAAAACAGC ATTCAGGTA
1201 TTAGAAGAAT ATCCTGATTC AGGTGAAAAT ATTGTTGATG CGCTGGCAGT GTTCTGCGC
1261 CGGTTGCATT CGATTCCTGT TTGTAATTGT CCTTTTAACA GCGATCGCGT ATTTCTGCTC
1321 GCTCAGGCGC AATCACGAAT GAATAACGGT TTGGTTGATG CGAGTGATTT TGATGACGAG
1381 CGTAATGGCT GGCCTGTTGA ACAAGTCTGG AAAGAAATGC ATAACTTTT GCCATTCTCA
1441 CCGGATTGAG TCGTCACTCA TGGTGATTTT TCACCTGATA ACCTTATTTT TGACGAGGGG
1501 AAATTAATAG GTTGTATTGA TGTGGACGA GTCGGAATCG CAGACCGATA TCAGGATCTT
1561 GCCATCCTAT GGAAGTGCCT CGGTGAGTTT TCTCCTTCAT TACAGAAACG GCTTTTTCAA
1621 AAATATGGTA TTGATAATCC TGATATGAAT AAATTGCAGT TTCATTTGAT GCTCGATGAG
1681 TTTTCTAAT CAGAATTGGT TAATTGGTTG TAACATTATT CAGATTGGGC CCCGTTCCAC
1741 TGAGCGTCAG ACCCGGTAGA AAAGATCAAA GGATCTTCTT GAGATCCTTT TTTTCTGCGC
1801 GTAATCTGCT GCTTGCAAAC AAAAAAACA CCGTACCAG CGGTGGTTT TTTGCCGGAT
1861 CAAGAGCTAC CAAGCTTTT TCCGAAGGTA ACTGGCTTCA GCAGAGCGCA GATACCAAAT
1921 ACTGTTCTTC TAGTGTAGCC GTAGTTAGGC CACCCTTCA AGAACTCTGT AGCACCCTCT
1981 ACATACCTCG CTCTGCTAAT CCTGTTACCA GTGGCTGCTG CCAGTGGCGA TAAGTCGTGT
2041 CTTACCGGGT TGGACTCAAG ACGATAGTTA CCGGATAAGG CGCAGCGGTC GGGCTGAACG
2101 GGGGGTTCGT GCACACAGCC CAGCTTGGAG CGAACGACCT ACACCGAACT GAGATACCTA
2161 CAGCGTGAGC TATGAGAAAG CGCCACGCTT CCCGAAGGGA GAAAGGCGGA CAGGTATCCG
2221 GTAAGCGGCA GGGTCGGAAC AGGAGAGCGC ACGAGGGAGC TTCCAGGGGG AAACGCCTGG
2281 TATCTTTATA GTCCTGTCGG GTTTCGCCAC CTCTGACTTG AGCGTCGATT TTTGTGATGC
2341 TCGTCAGGGG GCGCGAGCCT ATGGAATAAC GCCAGCAACG CGGCCTTTTT ACGGTTCTCTG
2401 GCCTTTTGCT GGCCTTTTGC TCACATGTTT TTTCTGCGT TATCCCCTGA TTCTGTGGAT
2461 AACCGTATTA CCGCTAGCAT GGATCTCGGG GACGTCTAAC TACTAAGCGA GAGTAGGGAA
2521 CTGCCAGGCA TCGAATAAAA CGAAAGGCTC AGTCGGAAGA CTGGGCCTTT CGTTTTATCT
2581 GTTGTGTTGTC GGTGAACGCT CTCCTGAGTA GGACAAATCC GCCGGGAGCG GATTTGAACG
2641 TTTGTGAAGCA ACGGCCCGGA GGGTGCGGG CAGGACGCCC GCCATAAACT GCCAGGCATC
2701 AAACCTAAGCA GAAGGCCATC

```

FIGURE 14B

20/240

Figure 1A: Cloning sites of the Entry Vector pENTR6

Int attL1 Sph I Kpn I Xmn I Sfi I

~~--- tgg tac aaa aaa gca ggc tgc atg cga acc aat tca gtc~~
~~--- ~~aac~~ ~~atg~~ ~~tgc~~ ttc cgt ccg acc taa gct tgg tta agt cag~~

Leu Tyr Lys Lys Ala Gly Cys Met Arg Thr Asn Ser Val

BamHI KpnI EpeRI EclRI

gac tgg atc cgg tac cga att cgc --- Death --- aga att cgc
 cgg acc tag gcp atg gct taa ggc --- (cod8) --- tct taa ggc

Asp Trp Ile Arg Tyr Arg Ile

Nat XhoI EclRI XbaI Int attL2

~~ggc cgc act cga gat atc tag acc cag ctt tgc tgc acc aag ---~~
~~ccg ggc tga gct cta tag atc tgg gtc gaa aga aca tgc ttc ---~~

21/240

pENTR6 2717 bp

<u>Location (Base Nos.)</u>	<u>Gene Encoded</u>
67..166	attL1
321..626	ccdB
655..754	attL2
877..1686	KmR
1791..2364	ori

```

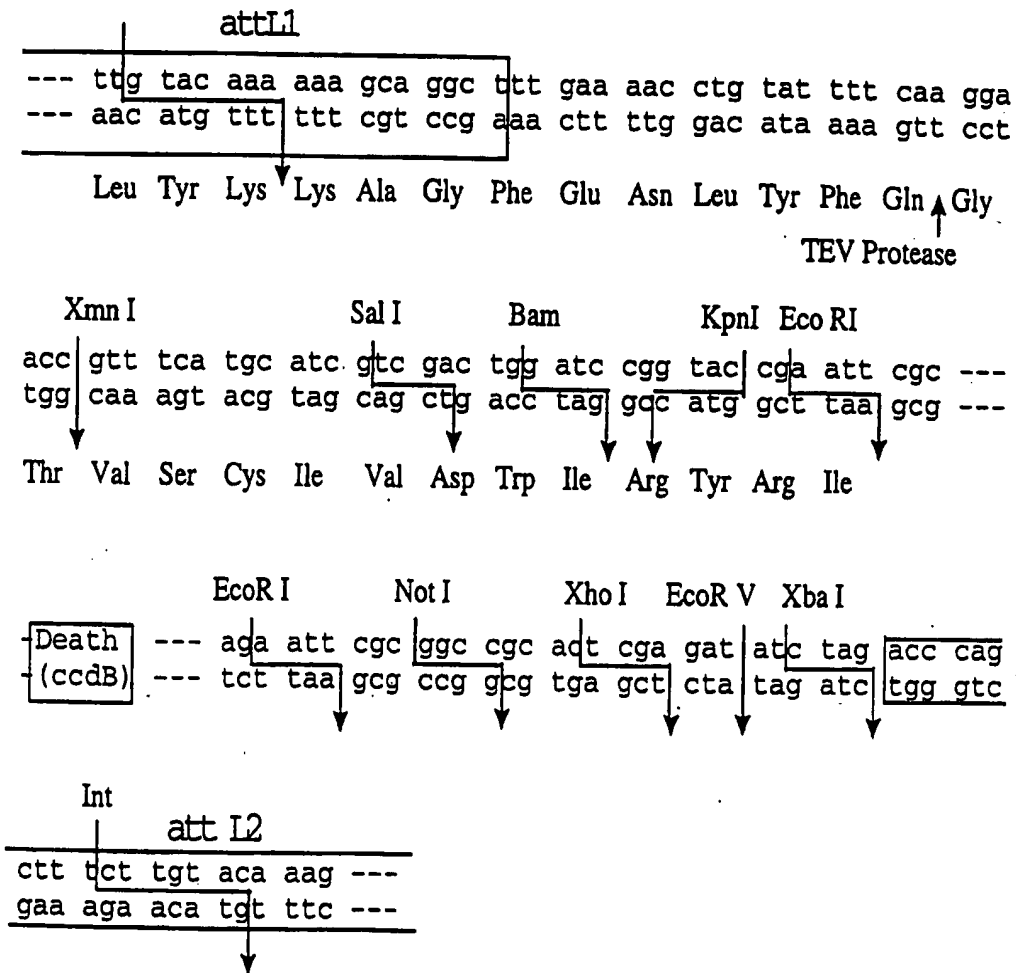
1 CTGACGGATG GCCTTTTTCG GTTTCTACAA ACTCTTCCTG TTAGTTAGTT ACTTAAGCTC
61 GGGCCCCAAA TAATGATTTT ATTTTGACTG ATAGTGACCT GTTCGTTGCA ACAAATTGAT
121 AAGCAATGCT TTTTATAAT GCCAACTTTG TACAAAAAAG CAGGCTGCAT GCGAACCAAT
181 TCAGTCGACT GGATCCGGTA CCGAATTCGC TTAATAAAG CCAGATAACA GTATGCGTAT
241 TTGCGCGCTG ATTTTTCGCG TATAAGAATA TATACTGATA TGTATACCCG AAGTATGTCA
301 AAAAGAGGTG TGCTTCTAGA ATGCAGTTTA AGGTTTACAC CTATAAAGA GAGAGCCGTT
361 ATCGTCTGTT TGTGGATGTA CAGAGTGATA TTATTGACAC GCCCGGGCGA CGGATGGTGA
421 TCCCCCTGGC CAGTGCACGT CTGCTGTCAG ATAAAGTCTC CCGTGAACCT TACCCGGTGG
481 TGCATATCGG GGATGAAAGC TGGCGCATGA TGACCACCGA TATGGCCAGT GTGCCGGTCT
541 CCGTTATCGG GGAAGAAGTG GCTGATCTCA GCCACCGCGA AAATGACATC AAAAACGCCA
601 TTAACCTGAT GTTCTGGGGA ATATAGAATT CGCGGCCGCA CTCGAGATAT CTAGACCCAG
661 CTTTCTTGTA CAAAGTTGGC ATTATAAGAA AGCATTGCTT ATCAATTGTT TGCAACGAAC
721 AGGTCACATAT CAGTCAAAAT AAAATCATT TTTGCCATCC AGCTGCAGCT CTGGCCCGTG
781 TCTCAAAATC TCTGATGTTA CATTGCACAA GATAAAATA TATCATCATG AACAAATAAA
841 CTGTCTGCTT ACATAAACAG TAATACAAGG GGTGTTATGA GCCATATTCA ACGGGAACG
901 TCGAGGCCGC GATTAAATTC CAACATGGAT GCTGATTAT ATGGGTATAA ATGGGCTCGC
961 GATAATGTCG GGCAATCAGG TGCACAATC TATCGCTTGT ATGGGAAGCC CGATGCGCCA
1021 GAGTTGTTTC TGAAACATGG CAAAGGTAGC GTTGCCAATG ATGTTACAGA TGAGATGGTC
1081 AGACTAAACT GGCTGACGGA ATTTATGCCT CTTCCGACCA TCAAGCATT TATCCGTAAT
1141 CCTGATGATG CATGGTTACT CACCACTGCG ATCCCCGAA AAACAGCATT CCAGGTATTA
1201 GAAGAATATC CTGATTCAGG TGAAAATATT GTTGATGCGC TGGCAGTGTT CCTGCGCCGG
1261 TTGCATTGTA TCCTGTTTG TAATTGTCCT TTTAACAGCG ATCGCGTATT TCGTCTCGCT
1321 CAGGCGCAAT CACGAATGAA TAACGGTTTG GTTGATGCGA GTGATTTTGA TGACGAGCGT
1381 AATGGCTGGC CTGTTGAACA AGTCTGAAA GAAATGCATA AACTTTTGCC ATTCTCACCG
1441 GATTAGTCTG TCACTCATGG TGATTCTCA CTTGATAACC TTATTTTGA CGAGGGGAAA
1501 TTAATAGGTT GTATTGATGT TGGACGAGTC GGAATCGCAG ACCGATACCA GGATCTTGGC
1561 ATCCTATGGA ACTGCCTCGG TGAGTTTCT CTTTCATTAC AGAAACGGCT TTTTCAAAAA
1621 TATGGTATTG ATAATCCTGA TATGAATAAA TTGCAGTTT ATTTGATGCT CGATGAGTTT
1681 TTCTAATCAG AATTGGTTAA TTGTTGTAA CATTATTCAG ATTGGGCCCC GTTCCACTGA
1741 GCGTCAGACC CCGTAGAAAA GATCAAAGGA TCTTCTTGAG ATCCTTTTTT TCTGCGCGTA
1801 ATCTGCTGCT TGCAAACAAA AAAACCACCG CTACCAGCGG TGGTTTGT TTGCCGATCAA
1861 GAGCTACCAA CTCTTTTTC GAAGGTAAT GGCTTCAGCA GAGCGCAGAT ACCAAATACT
1921 GTTCTTCTAG TGTAGCCGTA GTTAGGCCAC CACTTCAAGA ACTCTGTAGC ACCGCTTACA
1981 TACCTCGCTC TGCTAATCCT GTTACCAGTG GCTGTGCCA GTGGCGATAA GTCGTGCTT
2041 ACCGGGTTGG ACTCAAGACG ATAGTTACCG GATAAGGCGC AGCGGTCGGG CTGAACGGGG
2101 GGTTCGTGCA CACAGCCCAG CTTGGAGCGA ACGACCTACA CCGAAGTGA ATACCTACAG
2161 CGTGAGCTAT GAGAAAGCGC CACGCTTCCC GAAGGAGAAA AGGCGGACAG GTATCCGGTA
2221 AGCGGCAGGG TCGGAACAGG AGAGCGCAG AGGGAGCTTC CAGGGGGAAA CGCCTGGTAT
2281 CTTTATAGTC CTGTCCGGTT TCGCCACCTC TGACTTGAGC GTCGATTTT GTGATGCTCG
2341 TCAGGGGGGC GGAGCCTATG GAAAAACGCC AGCAACGCGG CCTTTTACG GTTCTTGGCC
2401 TTTTGCTGGC CTTTGTCTCA CATGTTCTTT CCTGCGTTAT CCCCTGATT TGTGGATAAC
2461 CGTATTACCG CTAGCATGGA TCTCGGGGAC GTCTAACTAC TAAGCGAGAG TAGGGAAC TG
2521 CCAGGCATCA AATAAAACGA AAGGCTCAGT CGGAAGACTG GGCCTTTCGT TTTATCTGTT
2581 GTTTGTGCGT GAACGCTCTC CTGAGTAGGA CAAATCCGCC GGGAGCGGAT TTGAACGTTG
2641 TGAAGCAACG GCCCGGAGGG TGGCGGGCAG GACGCCGCC ATAACTGCC AGGCATCAAA
2701 CTAAGCAGAA GGCCATC

```

Figure 15B

22/12/00

Figure 16A: Cloning sites of the ENTRY Vector PENTR7



pENTR7 2738 bp

<u>Location (Base No.)</u>	<u>Gene Encoded</u>
67..166	attL1
342..647	ccdB
676..775	attL2
898..1707	KmR
1812..2385	ori

```

1 CTGACGGATG GCCTTTTTCG GTTTCTACAA ACTCTTCCTG TTAGTTAGTT ACTTAAGCTC
61 GGGCCCCAAA TAATGATTTT ATTTTGACTG ATAGTGACCT GTTCGTTGCA ACAAATTGAT
121 AAGCAATGCT TTTTATAAT GCCAACTTTG TACAAAAAAG CAGGCTTTGA AAACCTGTAT
181 TTTCAAGGAA CCGTTTCATG CATCGTCGAC TGGATCCGGT ACCGAATTCG CTTACTAAAA
241 GCCAGATAAC AGTATGCGTA TTTGCGCGCT GATTTTTCG GTATAAGAAT ATATACTGAT
301 ATGTATACCC GAAGTATGTC AAAAAGAGGT GTGCTTCTAG AATGCAGTTT AAGGTTTACA
361 CCTATAAAG AGAGAGCCGT TATCGTCTGT TTGTGGATGT ACAGAGTGAT ATTATTGACA
421 CGCCCGGGCG ACGGATAGTG ATCCCCCTGG CCAGTGCACG TCTGCTGTCA GATAAAGTCT
481 CCCGTGAAC TACCCCGGTG GTGCATATCG GGGATGAAAG CTGGCGCATG ATGACCACCG
541 ATATGGCCAG TGTGCCGTC TCCGTTATCG GGGAAAGAAGT GGCTGATCTC AGCCACCCCG
601 AAAATGACAT CAAAAACGCC ATTAACCTGA TGTTCCTGGG AATATAGAAT TCGCGGCCGC
661 ACTCGAGATA TCTAGACCCA GCTTCTTGT ACAAAGTTGG CATTATAAGA AAGCATTGCT
721 TATCAATTTG TTGCAACGAA CAGGTCACTA TCAGTCAAAA TAAAATCATT ATTTGCCATC
781 CAGCTGCAGC TCTGGCCCGT GTCTCAAAAT CTCTGATGTT ACATTGCACA AGATAAAAAAT
841 ATATCATCAT GAACAATAAA ACTGTCTGCT TACATAAACA GTAATACAAG GGGTGTATG
901 AGCCATATTC AACGGGAAC GTGAGGCCG CGATTAAATT CCAACATGGA TGCTGATTTA
961 TATGGGTATA AATGGGCTCG CGATAATGTC GGGCAATCAG GTGCGACAAT CTATCGCTTG
1021 TATGGGAAGC CCGATGCGCC AGAGTTGTTT CTGAAACATG GCAAAGGTAG CGTTGCCAAT
1081 GATGTTACAG ATGAGATGGT CAGACTAAAC TGGCTGACGG AATTTATGCC TCTTCGACC
1141 ATCAAGCATT TTATCCGTAC TCCTGATGAT GCATGGTTAC TCACCACTGC GATCCCGGA
1201 AAAACAGCAT TCCAGGTATT AGAAGAATAT CCTGATTCAG GTGAAATAT TGTTGATGCG
1261 CTGGCAGTGT TCCTGCGCCG GTTGCAATTC ATTCTGTTT GTAATTGTCC TTTTAACAGC
1321 GATCGCGTAT TTCGTCTCGC TCAGGCGCAA TCACGAATGA ATAACGGTTT GGTTGATGCG
1381 AGTGATTTTG ATGACGAGCG TAATGGCTGG CCTGTTGAAC AAGTCTGGAA AGAAATGCAT
1441 AAACCTTTTC CATTCTCACC GGATTCAGTC GTCACTCATG GTGATTTCTC ACTTGATAAC
1501 CTTATTTTTC ACGAGGGGAA ATTAATAGGT TGTATTGATG TTGGACGAGT CGGAATCGCA
1561 GACCGATACC AGGATCTTGC CATCCTATGG AACTGCCTCG GTGAGTTTTC TCCTTCATTA
1621 CAGAAACGGC TTTTCAAAA ATATGGTATT GATAATCCTG ATATGAATAA ATTGCAGTTT
1681 CATTGATGTC TCGATGAGTT TTTCTAATCA GAATTGGTTA ATTGGTTGTA ACATTATTC
1741 GATTGGGCCC CGTTCCACTG AGCGTCAGAC CCCGTAGAAA AGATCAAAGG ATCTTCTTGA
1801 GATCCTTTTT TTCTGCGCGT AATCTGCTGC TTGCAACAA AAAAACCACC GCTACCAGCG
1861 GTGGTTTGTT TGCCGGATCA AGAGCTACCA ACTCTTTTTC CGAAGGTAAC TGGCTTCAGC
1921 AGAGCGCAGA TACCAATAC TGTCTTCTA GTGTAGCCGT AGTTAGGCCA CCACCTCAAG
1981 AACTCTGTAG CACCGCCTAC ATACCTCGCT CTGCTAATCC TGTACCAGT GGCTGCTGCC
2041 AGTGGCGATA AGTCGTGTCT TACCGGGTTG GACTCAAGAC GATAGTTACC GGATAAGGCG
2101 CAGCGGTCGG GCTGAACGGG GGGTTCGTGC ACACAGCCCA GCTTGGAGCG AACGACCTAC
2161 ACCGAACTGA GATACCTACA GCGTGAGCTA TGAGAAAGCG CCACGCTTCC CGAAGGGAGA
2221 AAGGCGGACA GGTATCCGGT AAGCGGCAGG GTCGGAACAG GAGAGCGCAT GAGGGAGCTT
2281 CCAGGGGGAA ACGCCTGGTA TCTTTATAGT CCGTTCGGGT TTCGCCACCT CTGACTTGAG
2341 CGTCGATTTT TGTGATGCTC GTCAGGGGGG CGGAGCCTAT GGAAAAACGC CAGCAACGCG
2401 GCCTTTTAC GGTTCCTGGC CTTTGTCTGG CCTTTTGCTC ACATGTTCTT TCCTGCGTTA
2461 TCCCCTGATT CTGTGGATAA CCGTATTACC GCTAGCATGG ATCTCGGGGA CGTCTAATA
2521 CTAAGCGAGA GTAGGGAACT GCCAGGCATC AAATAAAACG AAGGCTCAG TCGGAAGACT
2581 GGGCCTTTTC TTTTATCTGT TGTGTCGG TGAACGCTCT CCTGAGTAGG ACAAATCCGC
2641 CGGAGCGGA TTTGAACGTT GTGAAGCAAC GGCCCGGAGG GTGGCGGGCA GGACGCCCGC
2701 CATAAAGTGC CAGGCATCAA ACTAAGCAGA AGGCCATC

```

Figure 16B

24/240

Figure 17A: Cloning Sites of the Entry Vector pENTR8

Int attL

```

/---/ ttg tac aaa aaa gca ggc ttt gaa aac ctg tat ttt caa gga
/---/ acc atg atc ttt cgt ccg aaa cct ttg gac ata aaa gtt cct

```

Leu Tyr Lys Lys Ala Gly Phe Glu Asn Leu Tyr Phe Gln Gly

TEV Protease

NotI HaeII SalI BamHI KpnI EcoRI

```

acc atg gac cta gtc gac tgg atc cgg tac cga att cgc ---
tgg tac ctg gat cag cgg acc tag gct atg gct taa gcg ---

```

Thr Met Asp Leu Val Asp Trp Ile Arg Tyr Arg Ile

EcoRI NotI XhoI EcoRV XbaI attL

```

Death --- aga att cgc ggc cgc act cga gat atc tag acc cag
--- tct taa gcg ccg ggc tga gct cta tag atc tgg gtc

```

Int

```

cct tct atc aga tag ---
gaa aga aca tgt ttc ---

```

25/240

pENTR8 2735 bp

<u>Location (Base Nos.)</u>		<u>Gene Encoded</u>
67..166		attL1
339..644		ccdB
673..772		attL2
895..1704		KmR
1809..2382		ori
1	CTGACGGATG GCCTTTTTCG GTTCTACAA	ACTCTTCCTG TTAGTTAGTT ACTTAAGCTC
61	GGGCCCCAAA TAATGATTTT ATTTTGACTG	ATAGTGACCT GTTCGTTGCA ACAAATTGAT
121	AAGCAATGCT TTTTATAAT GCCAACTTTG	TACAAAAAAG CAGGCTTTGA AAACCTGTAT
181	TTTCAAGGAA CCATGGACCT AGTCGACTGG	ATCCGGTACC GAATTCGCTT ACTAAAAGCC
241	AGATAACAGT ATGCGTATTT GCGCGCTGAT	TTTTGCGGTA TAAGAATATA TACTGATATG
301	TATACCCGAA GTATGTCAAA AAGAGGTGTG	CTTCTAGAAT GCAGTTTAAG GTTTACACCT
361	ATAAAAGAGA GAGCCGTTAT CGTCTGTTTG	TGGATGTACA GAGTGATATT ATTGACACGC
421	CCGGGCGACG GATAGTGATC CCCTGGCCA	GTGCACGTCT GCTGTCAGAT AAAGTCTCCC
481	GTGAACCTTA CCCGGTGGTG CATATCGGGG	ATGAAAGCTG GCGCATGATG ACCACCGATA
541	TGGCCAGTGT GCCGGTCTCC GTTATCGGGG	AAGAAGTGGC TGATCTCAGC CACCGCGAAA
601	ATGACATCAA AAACGCCATT AACCTGATGT	TCTGGGGAAT ATAGAATTCTG CGGCCGCACT
661	CGAGATATCT AGACCCAGCT TTCTTGATCA	AAGTTGGCAT TATAAGAAAG CATTGCTTAT
721	CAATTTGTTG CAACGAACAG GTCACATATCA	GTCAAATAA AATCATTATT TGCCATCCAG
781	CTGCAGCTCT GCGCCGTGTC TCAAAATCTC	TGATGTTACA TTGCACAAGA TAAAAATATA
841	TCATCATGAA CAATAAACT GTCTGCTTAC	ATAAACAGTA ATACAAGGGG TGTTATGAGC
901	CATATTCAAC GGGAAACGTC GAGGCCGCGA	TTAAATTCCA ACATGGATGC TGATTTATAT
961	GGGTATAAAT GGGCTCGCGA TAATGTCGGG	CAATCAGGTG CGACAATCTA TCGCTTGAT
1021	GGGAAGCCCG ATGCGCCAGA GTTGTTCCTG	AAACATGGCA AAGGTAGCGT TGCCAAATGAT
1081	GTTACAGATG AGATGGTCAG ACTAACTGG	CTGACGGAAT TTATGCCTCT TCCGACCATC
1141	AAGCATTTTA TCCGTACTCC TGATGATGCA	TGGTTACTCA CCACTGCGAT CCCCGGAAAA
1201	ACAGCATTCC AGGTATTAGA AGAATATCCT	GATTCAGGTG AAAATATTGT TGATGCGCTG
1261	GCAGTGCTCC TGCGCCGTT GCATTGATG	CCTGTTTGTA ATTGTCCTTT TAACAGCGAT
1321	CGCGTATTTT GTCTCGCTCA GCGCAATCA	CGAATGAATA ACGGTTTGGT TGATGCGAGT
1381	GATTTTGATG ACGAGCGTAA TGCTGGCCT	GTTGAACAAG TCTGGAAAGA AATGCATAAA
1441	CTTTTGCCAT TCTACCCGGA TTCAGTCGTC	ACTCATGGTG ATTTCTCACT TGATAACCTT
1501	ATTTTGTGAC AGGGGAAATT AATAGGTTGT	ATTGATGTTG GACGAGTCGG AATCGCAGAC
1561	CGATACCAGG ATCTTGCCAT CCTATGGAAC	TGCCTCGGTG AGTTTCTCTC TTCATTACAG
1621	AAACGGCTTT TTCAAAAATA TGGTATTGAT	AATCCTGATA TGAATAAATT GCAGTTTCAT
1681	TTGATGCTCG ATGAGTTTCT CTAATCAGAA	TTGGTTAATT GGTGTAACA TTATTAGAT
1741	TGGGCCCCGT TCCACTGAGC GTCAGACCCC	GTAAGAAAGA TCAAAGGATC TTCTTGAGAT
1801	CCTTTTTCCT TGCGCGTAAT CTGCTGCTTG	CAAAACAAAA AACCACCGCT ACCAGCGGTG
1861	GTTTGTTTGC CGGATCAAGA GCTACCAACT	CTTTTCCGA AGGTAAGTGG CTTCAGCAGA
1921	GCGCAGATAC CAAATACTGT TCTTCTAGTG	TAGCCGTAGT TAGGCCACCA CTTCAGAAC
1981	TCTGTAGCAC CGCCTACATA CCTCGCTCTG	CTAATCCTGT TACCAGTGGC TGCTGCCAGT
2041	GGCGATAAGT CGTGTCTTAC CGGGTTGGAC	TCAAGACGAT AGTTACCGGA TAAGGCGCAG
2101	CGTTCGGGCT GAACGGGGGG TTCGTGCACA	CAGCCCAGCT TGGAGCGAAC GACCTACACC
2161	GAACTGAGAT ACCTACAGCG TGAGCTATGA	GAAAGCGCCA CGCTTCCCGA AGGGAGAAAAG
2221	GCGGACAGGT ATCCGGTAAG CGGCAGGGTC	GGAACAGGAG AGCGCACGAG GGAGCTTCCA
2281	GGGGGAAACG CCTGGTATCT TTATAGTCCT	GTCGGGTTTC GCCACCTCTG ACTTGAGCGT
2341	CGATTTTGTG GATGCTCGTC AGGGGGGCGG	AGCCTATGGA AAAACGCCAG CAACGCGGCC
2401	TTTTTACGGT TCCTGGCCTT TTGCTGGCCT	TTTGCTCACA TGTTCTTTCC TGCGTTATCC
2461	CCTGATTCTG TGGATAACCG TATTACCGCT	AGCATGGATC TCGGGGACGT CTAATACTA
2521	AGCGAGAGTA GGGAACTGCC AGGCATCAAA	TAAAACGAAA GGCTCAGTCG GAAGACTGGG
2581	CCTTTCGTTT TATCTGTTGT TTGTCGGTGA	ACGCTCTCCT GAGTAGGACA AATCCGCGCG
2641	GAGCGGATTT GAACGTTGTG AAGCAACGGC	CCGGAGGGTG GCGGGCAGGA CGCCCGCCAT
2701	AAACTGCCAG GCATCAAACT AAGCAGAAGG	CCATC

FIGURE 17B

26/240

Figure 18A: Cloning sites of the ENTRY Vector pENTRY

~~Int~~ ~~attL1~~
~~--- /~~ ~~ctg~~ ~~tac~~ ~~aaa~~ ~~aaa~~ ~~gca~~ ~~ggc~~ ~~ttt~~ ~~gaa~~ ~~aac~~ ~~ctg~~ ~~tat~~ ~~ttt~~ ~~caa~~ ~~gga~~
~~--- /~~ ~~aaq~~ ~~gag~~ ~~ctc~~ ~~ttt~~ ~~cgt~~ ~~ccg~~ ~~aaa~~ ~~ctt~~ ~~ttg~~ ~~gac~~ ~~ata~~ ~~aaa~~ ~~gtt~~ ~~cct~~
 Leu Tyr Lys Lys Ala Gly Phe Glu Asn Leu Tyr Phe Gln Gly
 TEV protease

NdeI BglII SalI BamHI KpnI EcoRI
 cat atg aga tct gtc gac tgc atc cgg tac cga att cgc ---
 gta tgc tct aga cag ctg acc tag ggc atg gct taa gcg ---
 His Met Arg Ser Val Asp Trp Ile Arg Tyr Arg Ile

EcoRI NotI XhoI EcoRI XbaI attL2
 Death --- aga att cgc ggc cgc act cga gat atc tag acc cag
 --- tct taa gcg ccg gcg tga gct cta tag atc tgg gtc

Int
 ctt ~~ctt~~ ~~tgt~~ ~~aca~~ ~~tag~~ ---
 gaa aga aca tct ~~tcc~~ ---

27/240

pENTR9 2735 bp

<u>Location (Base Nos.)</u>		<u>Gene Encoded</u>
67..166		attL1
339..644		ccdB
673..772		attL2
895..1704		KmR
1809..2382		ori
1	CTGACGGATG GCCTTTTTGC GTTTCTACAA ACTCTTCCTG TTAGTTAGTT ACTTAAGCTC	
61	GGGCCCCAAA TAATGATTTT ATTTTGACTG ATAGTGACCT GTTCGTTGCA ACAAATTGAT	
121	AAGCAATGCT TTTTATAAT GCCAATTG TACAAAAAAG CAGGCTTTGA AAACCTGTAT	
181	TTTCAAGGAC ATATGAGATC TGTCGACTGG ATCCGGTACC GAATTCGCTT ACTAAAAGCC	
241	AGATAACAGT ATGCGTATTT GCGCGCTGAT TTTTGCGGTA TAAGAATATA TACTGATATG	
301	TATACCCGAA GTATGTCAAA AAGAGGTGTG CTCTAGAAT GCAGTTTAAG GTTTACACCT	
361	ATAAAAGAGA GAGCCGTTAT CGTCTGTTTG TGGATGTACA GAGTGATATT ATTGACACGC	
421	CCGGGCGACG GATAGTGATC CCCCTGGCCA GTGCACGTCT GCTGTCAGAT AAAGTCTCCC	
481	GTGAACCTTA CCCGGTGGTG CATATCGGGG ATGAAAGCTG GCGCATGATG ACCACCGATA	
541	TGGCCAGTGT GCCGGTCTCC GTTATCGGGG AAGAAGTGGC TGATCTCAGC CACCGCGAAA	
601	ATGACATCAA AAACGCCATT AACCTGATGT TCTGGGGAAT ATAGAATTCC CGGCCGCACT	
661	CGAGATATCT AGACCCAGCT TTCTGTACA AAGTTGGCAT TATAAGAAAG CATTGCTTAT	
721	CAATTTGTTG CAACGAACAG GTCACATCA GTCAAAATAA AATCATTATT TGCCATCCAG	
781	CTGCAGCTCT GGCCCGTGTC TCAAAATCTC TGATGTTACA TTGCACAAGA TAAAAATATA	
841	TCATCATGAA CAATAAACT GTCTGCTTAC ATAAACAGTA ATACAAGGGG TGTTATGAGC	
901	CATATTCAAC GGGAAACGTC GAGGCCGCGA TTAAATTCCA ACATGGATGC TGATTTATAT	
961	GGGTATAAAT GGGCTCGCGA TAATGTCGGG CAATCAGGTG CGACAATCTA TCGCTTGTAT	
1021	GGGAAGCCCG ATGCGCCAGA GTTGTTTCTG AAACATGGCA AAGGTAGCGT TGCCAATGAT	
1081	GTTACAGATG AGATGGTCAG ACTAACTGG CTGACGGAAT TTATGCCTCT TCCGACCATC	
1141	AAGCATTTTA TCCGTACTCC TGATGATGCA TGGTTACTCA CCACTGCGAT CCCCAGAAAA	
1201	ACAGCATTTCC AGGTATTAGA AGAATATCCT GATTCAGGTG AAAATATTGT TGATGCGCTG	
1261	GCAGTGTCCTG TCGCGCGGTT GCATTTCGATT CCTGTTTGTA ATTGTCCTTT TAACAGCGAT	
1321	CGCGTATTTT GTCTCGCTCA GCGCAATCA CGAATGAATA ACGGTTTGGT TGATGCGAGT	
1381	GATTTTGATG ACGAGCGTAA TGGCTGGCCT GTTGAACAAG TCTGGAAAGA AATGCATAAA	
1441	CTTTTGCCAT TCTCACCGGA TTCAGTCGTC ACTCATGGTG ATTTCTCACT TGATAACCTT	
1501	ATTTTGTACG AGGGGAAATT AATAGTTGT ATTGATGTTG GACGAGTCGG AATCGCAGAC	
1561	CGATACCAGG ATCTTGCCAT CCTATGGAAC TGCCCTCGGTG AGTTTTCTCC TTCATTACAG	
1621	AAACGGCTTT TTCAAAAATA TGGTATTGAT AATCCTGATA TGAATAAATT GCAGTTTCAT	
1681	TTGATGCTCG ATGAGTTTTT CTAATCAGAA TTGGTTAATT GGTGTGAACA TTATTCAGAT	
1741	TGGGCCCCGT TCCACTGAGC GTCAGACCCC GTAGAAAAGA TCAAAGGATC TTCTTGAGAT	
1801	CCTTTTTCCTG TCGCGTAAT CTGCTGCTTG CAAACAAAAA AACCACCGCT ACCAGCGGTG	
1861	GTTTGTTTGC CGGATCAAGA GCTACCAACT CTTTTTCCGA AGGTAACCTG CTTACGCAGA	
1921	GCGCAGATAC CAAATACTGT TCTTCTAGTG TAGCCGTAGT TAGGCCACCA CTTCAAGAAC	
1981	TCTGTAGCAC CGCCTACATA CCTCGCTCTG CTAATCCTGT TACCAGTGGC TGCTGCCAGT	
2041	GGCGATAAGT CGTGTCTTAC CGGGTTGGAC TCAAGACGAT AGTTACCGGA TAAGGCGCAG	
2101	CGGTCGGGCT GAACGGGGGG TTCGTGCACA CAGCCCAGCT TGGAGCGAAC GACCTACACC	
2161	GAAGTGAAGT ACCTACAGCG TGAGCTATGA GAAAGCGCCA CGCTTCCCGA AGGGAGAAAG	
2221	GCGGACAGGT ATCCGGTAAG CGGCAGGGTC GGAACAGGAG AGCGCACGAG GAGCTTCCCA	
2281	GGGGGAAACG CCTGGTATCT TTATAGTCCT GTCGGGTTTC GCCACCTCTG ACTTGAGCGT	
2341	CGAATTTTGT GATGCTCGTC AGGGGGGCGG AGCCTATGGA AAAACGCCAG CAACGCGGCC	
2401	TTTTTACGGT TCCTGGCCTT TTGCTGGCCT TTTGCTCACA TGTTCTTTCC TGCGTTATCC	
2461	CCTGATTCTG TGGATAACCG TATTACCGCT AGCATGGATC TCGGGGACGT CTAACACTA	
2521	AGCGAGAGTA GGGAACTGCC AGGCATCAAA TAAAACGAAA GGCTCAGTCG GAAGACTGGG	
2581	CCTTTCGTTT TATCTGTTGT TTGTCGGTGA ACGCTCTCCT GAGTAGGACA AATCCGCCGG	
2641	GAGCGGATTT GAACGTTGTG AAGCAACGGC CCGGAGGGTG GCGGGCAGGA CGCCCGCCAT	
2701	AAACTGCCAG GCATCAAACCT AAGCAGAAGG CCATC	

FIGURE 18B

28/240

Figure 19A: Cloning sites of the ENTRY Vector pENTR10

Int attL1 S.D. -12 Hek

--- ~~ctg tac aaa aaa gca ggc ttc~~ gaa cta agg aaa tac tta cgt
 --- ~~aac atg ttc~~ ttt cgt ccg aag ctt gat tcc ttt atg aat gta
 Leu Tyr Lys Lys Ala Gly Phe Glu Leu Arg Lys Tyr Leu His

K3 Xon Sal Bam Kpn EcoRI

atg gga acc aat tca gtc gac tgg atc cgg tac cga att cgc ---
 tac cct tgg tta agt cag cgg acc tag gop atg gct taa gog ---
 Met Gly Thr Asn Ser Val Asp Trp Ile Arg Tyr Arg Ile

EcoRI Not Xho EcoRII Xba attL2

Death --- aga att cgc ggc cgc act cga gat atc tag acc cag
 (ccdB) --- tct aaa gcg ccg ggc tga gct cta tag atc tgg gtc

Int

ctt ttt tgg aca aag
gaa aga aca tga ttc

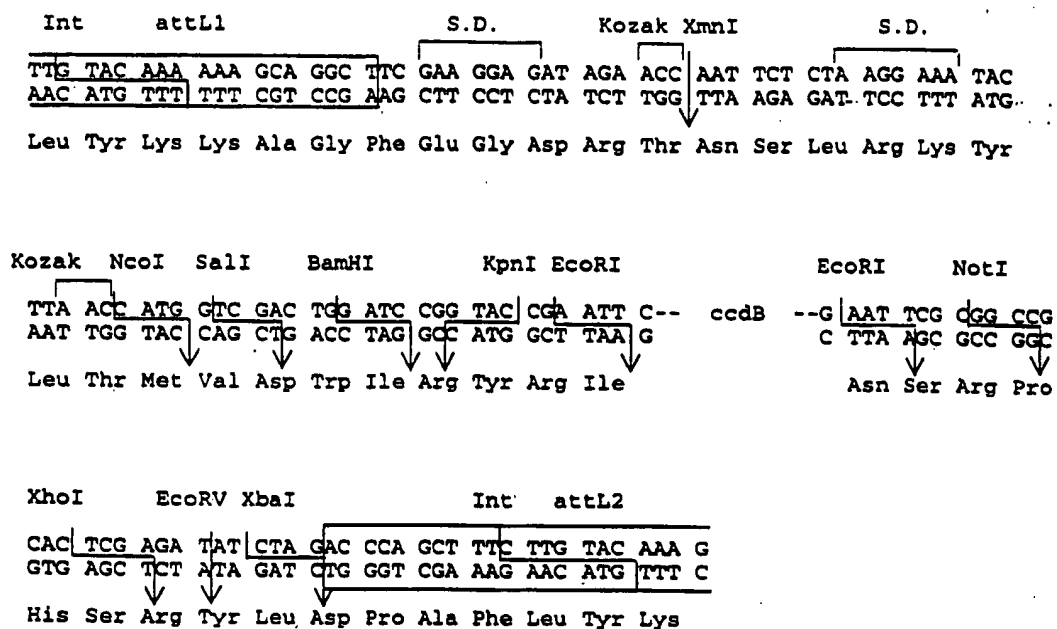
pENTR10 2738 bp

<u>Location (Base Nos.)</u>		<u>Gene Encoded</u>
67..166		attL1
342..647		ccdB
676..775		attL2
898..1707		KmR
1812..2385		ori

1	CTGACGGATG	GCCTTTTTCG	GTTTCTACAA	ACTCTTCCTG	TTAGTTAGTT	ACTTAAGCTC
61	GGGCCCCAAA	TAATGATTTT	ATTTTGACTG	ATAGTGACCT	GTCGTTGCA	ACAAATGAT
121	AAGCAATGCT	TTTTTATAAT	GCCAACTTTG	TACAAAAAAG	CAGGCTTCGA	ACTAAGGAAA
181	TACTTACATA	TGGGAACCAA	TTCAGTCGAC	TGGATCCGGT	ACCGAATTCG	CTTACTAAAA
241	GCCAGATAAC	AGTATGCGTA	TTTGC GCGCT	GATTTTTCG	GTATAAGAAT	ATATACTGAT
301	ATGTATACCC	GAAGTATGTC	AAAAAGAGGT	GTGCTTCTAG	AATGCAGTTT	AAGGTTTACA
361	CCTATAAAAG	AGAGAGCCGT	TATCGTCTGT	TTGTGGATGT	ACAGAGTGAT	ATTATTGACA
421	CGCCCGGGCG	ACGGATGGTG	ATCCCCCTGG	CCAGTGCACG	TCTGCTGTCA	GATAAAGTCT
481	CCCGTGAACT	TTACCCGGTG	GTGCATATCG	GGGATGAAAG	CTGGCGCATG	ATGACCACCG
541	ATATGGCCAG	TGTGCCGGTC	TCCGTTATCG	GGGAAGAAAGT	GGCTGATCTC	AGCCACCGCG
601	AAAATGACAT	CAAAAACGCC	ATTAACCTGA	TGTTCTGGGG	AATATAGAAT	TCGCGGCCGC
661	ACTCGAGATA	TCTAGACCCA	GCTTTCTTGT	ACAAAGTTGG	CATTATAAGA	AAGCATTGCT
721	TATCAATTTG	TTGCAACGAA	CAGGTCACCTA	TCAGTCAAAA	TAAATCATT	ATTTGCCATC
781	CAGCTGCAGC	TCTGGCCCGT	GTCTCAAAAT	CTCTGATGTT	ACATTGCACA	AGATAAAAAAT
841	ATATCATCAT	GAACRAATAA	ACTGTCTGCT	TACATAAACA	GTAATACAAG	GGGTGTTATG
901	AGCCATATTC	AACGGGAAAC	GTGAGGGCCG	CGATTAAATT	CCAACATGGA	TGCTGATTTA
961	TATGGGTATA	AATGGGCTCG	CGATAATGTC	GGGCAATCAG	GTGCGACAAT	CTATCGCTTG
1021	TATGGGAAGC	CCGATGCGCC	AGAGTTGTTT	CTGAAACATG	GCAAAGGTAG	CGTTGCCAAT
1081	GATGTTACAG	ATGAGATGGT	CAGACTAAAC	TGGCTGACGG	AATTTATGCC	TCTTCCGACC
1141	ATCAAGCATT	TTATCCGTAC	TCCTGATGAT	GCATGGTTAC	TCACCACTGC	GATCCCCGGA
1201	AAAACAGCAT	TCCAGGTATT	AGAAGAATAT	CCTGATTCAG	GTGAAAATAT	TGTTGATGCG
1261	CTGGCAGTGT	TCCTGCGCCG	GTTGCATTTC	ATTCTGTGTT	GTAATTGTCC	TTTTAACAGC
1321	GATCGCGTAT	TTCTCTCTCG	TCAGGCGCAA	TCACGAATGA	ATAACGGTTT	GGTTGATGCG
1381	AGTGATTTTG	ATGACGAGCG	TAATGGCTGG	CCTGTTGAAC	AAGTCTGGAA	AGAAATGCAT
1441	AAACTTTTTC	CATTCTCACC	GGATTTCAGT	GTCACTCATG	GTGATTTCTC	ACTTGATAAC
1501	CTTATTTTTC	ACGAGGGGAA	ATTAATAGGT	TGTATTGATG	TTGGACGAGT	CGGAATCGCA
1561	GACCGATACC	AGGATCTTGC	CATCCTATGG	AACTGCCTCG	GTGAGTTTTC	TCCTTCATTA
1621	CAGAAACGGC	TTTTTCAAAA	ATATGGTATT	GATAATCTCG	ATATGAATAA	ATTGCACTTT
1681	CATTTGATGC	TCGATGAGTT	TTTCTAATCA	GAATTGGTTA	ATTGGTTGTA	ACATTATTCA
1741	GATTGGGCCC	CGTTCCACTG	AGCGTCAGAC	CCCGTAGAAA	AGATCAAAGG	ATCTTCTTGA
1801	GATCCTTTTT	TTCTGCGCGT	AATCTGCTGC	TTGCAAACAA	AAAAACCACC	GCTACCAGCG
1861	GTGGTTTGTT	TGCCGGATCA	AGAGCTACCA	ACTCTTTTTC	CGAAGGTAAC	TGGCTTCAGC
1921	AGAGCGCAGA	TACCAATATC	TGTTCTTCTA	GTGTAGCCGT	AGTTAGGCCA	CCACTTCAAG
1981	AACTCTGTAG	CACCGCTTAC	ATACCTCGCT	CTGCTAATCC	TGTTACCAGT	GCTGCTGCC
2041	AGTGGCGATA	AGTCGTGTCT	TACCGGGTTG	GACTCAAGAC	GATAGTTACC	GGATAAGGCG
2101	CAGCGGTCGG	GCTGAACGGG	GGGTTCTGTC	ACACAGCCCA	GCTTGAGCG	AACGACCTAC
2161	ACCGAACTGA	GATACCTACA	GCGTGAGCTA	TGAGAAAGCG	CCACGCTTCC	CGAAGGGAGA
2221	AAGGCGGACA	GGTATCCGGT	AAGCGGCAGG	GTCGGAACAG	GAGAGCGCAC	GAGGGAGCTT
2281	CCAGGGGGAA	ACGCCTGGTA	TCTTTATAGT	CCTGTCCGGT	TTCCGCCACT	CTGACTTGAG
2341	CGTCGATTTT	TGTGATGCTC	GTGAGGGGGG	CGGAGCCTAT	GGAAAAACGC	CAGCAACGCG
2401	GCTTTTTC	GGTTCCCTGG	CTTTTGCTGG	CCTTTTGCTC	ACATGTTCTT	TCCTGCGTTA
2461	TCCCTTGATT	CTGTGGATAA	CCGTATTACC	GCTAGCATGG	ATCTCGGGGA	CGTCTAATA
2521	CTAAGCGAGA	GTAGGGAACT	GCCAGGCATC	GAATAAAACG	AAAGGCTCAG	TCGGAAGACT
2581	GGGCCTTTTC	TTTTATCTGT	TGTTTGTCGG	TGAACGCTCT	CCTGAGTAGG	ACAAATCCGC
2641	CGGGAGCGGA	TTTGAACGTT	GTGAAGCAAC	GGCCCGGAGG	GTGGCGGGCA	GGACGCCCCG
2701	CATAAACTGC	CAGGCATCAA	ACTAAGCAGA	AGGCCATC		

FIGURE 19B

30/240

Figure 20A: Cloning Sites of the Entry Vector pENTR11

31/240

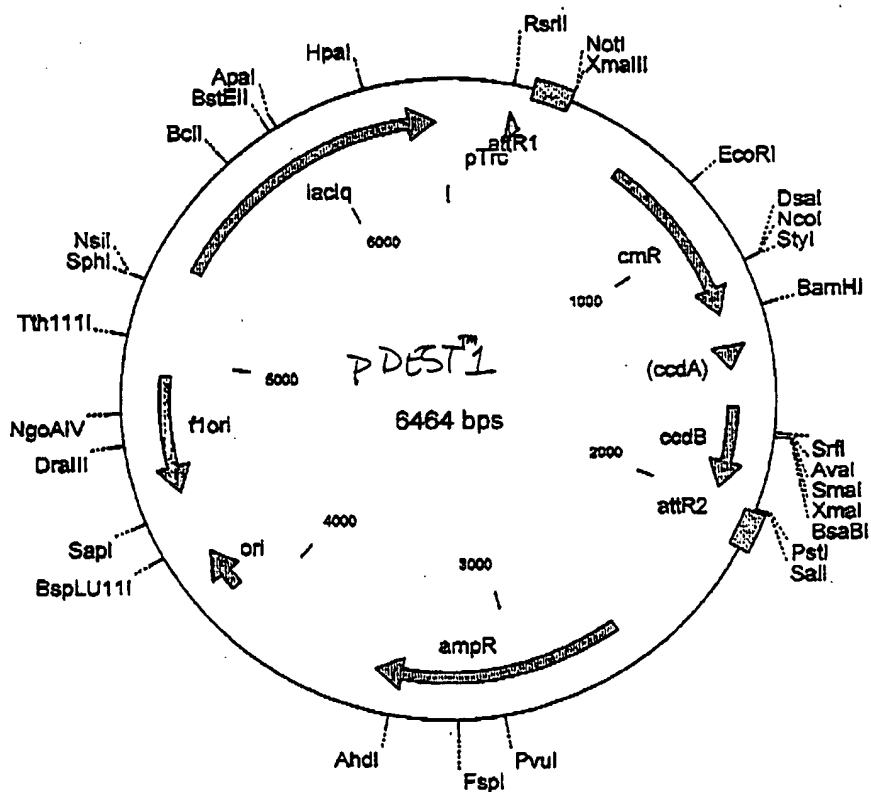
pENTR11 2744 bp (rotated to position 2578)

<u>Location (Base Nos.)</u>		<u>Gene Encoded</u>
67..166		attL1
348..653		ccdB
683..781		attL2
904..1713		KmR
1818..2391		ori
1	CTGACGGATG GCCTTTTTCG GTTCTACAA ACTCTTCCTG TTAGTTAGTT ACTTAAGCTC	
61	GGGCCCCAAA TAATGATTTT ATTTTGACTG ATAGTGACCT GTTCGTTGCA ACAAATTGAT	
121	AAGCAATGCT TTTTATAAT GCCAACTTTG TACAAAAAAG CAGGCTTCGA AGGAGATAGA	
181	ACCAATTCTC TAAGGAAATA CTTAACCATG GTCGACTGGA TCCGGTACCG AATTTCGCTTA	
241	CTAAAAGCCA GATAACAGTA TGCCTATTG CGCGCTGATT TTTGCGGTAT AAGAATATAT	
301	ACTGATATGT ATACCCGAAG TATGTCAAAA AGAGGTGTGC TTCTAGAATG CAGTTTAAGG	
361	TTTACACCTA TAAAAGAGAG AGCCGTTATC GTCTGTTTGT GGATGTACAG AGTGATATTA	
421	TTGACACGCC CGGGCGACGG ATAGTGATCC CCTGGCCAG TGCACGTCTG CTGTCAGATA	
481	AAGTCTCCCG TGAACCTTAC CCGGTGGTGC ATATCGGGGA TGAAAGCTGG GCATGTATGA	
541	CCACCGATAT GGCCAGTGTG CCGGTCTCCG TTATCGGGGA AGAAGTGGCT CATGTCAGCC	
601	ACCGCGAAAA TGACATCAAA AACGCCATTA ACCTGATGTT CTGGGGAATA TAGAATTCGC	
661	GGCCGCACTC GAGATATCTA GACCCAGCTT TCTGTACAA AGTTGGCATT ATAAGAAAGC	
721	ATTGCTTATC AATTGTGTG AACGAACAGG TCACTATCAG TCAAAATAAA ATCATTATTT	
781	GCCATCCAGC TGCAGCTCTG GCGGTGTCT CAAAATCTCT GATGTTACAT TGCACAAGAT	
841	AAAAATATAT CATCATGAAC AATAAACTG TCTGCTTACA TAAACAGTAA TACAAGGGGT	
901	GTTATGAGCC ATATTCAACG GGAACCTCG AGGCCGCGAT TAAATTCCAA CATGGATGCT	
961	GATTTATATG GGTATAAATG GGCTCGCGAT AATGTCGGGC AATCAGGTGC GACAATCTAT	
1021	CGCTTGATG GGAAGCCCGA TGCGCCAGAG TTGTTTCTGA AACATGGCAA AGGTAGCGTT	
1081	GCCAATGATG TTACAGATGA GATGGTCAGA CTAAACTGGC TGACCGAATT TATGCCTCTT	
1141	CCGACCATCA AGCATTTTAT CCGTACTCCT GATGATGCAT GGTACTCAC CACTGCGATC	
1201	CCCGGAAAAA CAGCATTCCA GGTATTAGAA GAATATCCTG ATTCAGGTGA AAATATTGTT	
1261	GATGCGCTGG CAGTGTTCC TCGCCGGTTG CATTCGATTG CTGTTTGTA TGTCTCTTTT	
1321	AACAGCGATC GCGTATTTG TCTCGCTCAG GCGCAATCAC GAATGAATAA CCGTTTGTTT	
1381	GATGCGAGTG ATTTTGATGA CGAGCGTAAT GGCTGGCCTG TTGAACAAGT CTGGAAAGAA	
1441	ATGCATAAAC TTTTGCCATT CTCACCGGAT TCAGTCGTCA CTCATGGTGA TTTCTCACTT	
1501	GATAACCTTA TTTTGTACGA GGGGAAATTA ATAGGTTGTA TTGATGTTGG ACGAGTCGGA	
1561	ATCGCAGACC GATACCAGGA TCTTGCCATC CTATGGAAT GCCTCGGTGA GTTTTCTCCT	
1621	TCATTACAGA AACGGCTTTT TCAAAAATAT GGTATTGATA ATCCTGATAT GAATAAATTG	
1681	CAGTTTCATT TGATGCTCGA TGAGTTTTC TAATCAGAAT TGGTTAATTG GTTGTAACAT	
1741	TATTCAGATT GGGCCCCGTT CCACTGAGCG TCAGACCCCG TAGAAAAGAT CAAAGGATCT	
1801	TCTTGAGATC CTTTTTTTCT GCGCGTAATC TGCTGCTTGC AAACAAAAAA ACCACCGCTA	
1861	CCAGCGGTGG TTTGTTTGCC GGATCAAGAG CTACCAACTC TTTTTCGGAA GGTAAGTGGC	
1921	TTCAGCAGAG CGCAGATACC AAATACTGTT CTCTAGTGT AGCCGTAGTT AGGCCACCAAC	
1981	TTCAAGAACT CTGTAGCACC GCCTACATAC CTCGCTCTGC TAATCCTGTT ACCAGTGGCT	
2041	GCTGCCAGTG GCGATAAGTC GTGTCTTACC GGGTTGGACT CAAGACGATA GTTACCGGAT	
2101	AAGGCGCAGC GGTGCGGCTG AACGGGGGGT TCGTGCACAC AGCCCGCTT GGAGCGAACG	
2161	ACCTACACCG AACTGAGATA CCTACAGCGT GAGCTATGAG AAAGCGCCAC GCTTCCCGAA	
2221	GGGAGAAAGG CGGACAGGTA TCCGGTAAGC GGCAGGGTCG GAACAGGAGA GCGCACGAGG	
2281	GAGCTTCCAG GGGGAAACGC CTGGTATCTT TATAGTCCTG TCGGGTTTCG CCACCTCTGA	
2341	CTTGAGCGTC GATTTTGTG ATGCTCGTCA GGGGGGCGGA GCCTATGGAA AAACGCCAGC	
2401	AACGCGCCT TTTACGGTT CCTGGCCTTT TGCTGGCCTT TTGCTCACAT GTTCTTTCTT	
2461	GCGTTATCCC CTGATTCTGT GGATAACCGT ATTACCGCTA GCATGGATCT CGGGGACGTC	
2521	TAATACTAA GCGAGAGTAG GGAAGTCCA GGCATCAAT AAAACGAAAG GCTCAGTCGG	
2581	AAGACTGGGC CTTTCGTTTT ATCTGTTGTT TGTCGGTGAA CGCTCTCCTG AGTAGGACAA	
2641	ATCCGCGGGG AGCGGATTTG AACGTTGTGA AGCAACGGCC CGGAGGGTGG CGGGCAGGAC	
2701	GCCCGCCATA AACTGCCAGG CATCAACTA AGCAGAAGGC CATC	

FIGURE 20B

Figure 2/A: *PDEST1* Native Protein Expression in *E. coli*

-35 Trc promoter -10 P_{MANA}
 1 atgagctgtt gacatttaant catccggctc gataaattg tggatttgg agcggataac
 tactcgacha ctgttaatta gtaggccgag catattacac accttaacac tcgectattg
 61 aatttcacac aggaacacaga caggtatagg atcacaagtt gtatgaada agctgaar
 ttaaagtgtg tcctttgtct gtccatatcc taggttcaa acatgtttc tcgacttgc



33/240

pDEST1 6464 bp

<u>Location (Base Nos.)</u>		<u>Gene Encoded</u>
216..257		Trc promoter
397..273		attR1
647..1306		CmR
1426..1510		inactivated ccdA
1648..1953		ccdB
1994..2118		attR2
2598..3503		ampR
4104..4264		ori
4504..4941		flori (f1 intergenic region)
5340..6420		lacIq

1	GTTTGACAGC	TTATCATCGA	CTGCACGGTG	CACCAATGCT	TCTGGCGTCA	GGCAGCCATC
61	GGAAGCTGTG	GTATGGCTGT	GCAGGTCGTA	AATCACTGCA	TAATTCGTGT	CGCTCAAGGC
121	GCACTCCCGT	TCTGGATAAT	GTTTTTTGCG	CCGACATCAT	AACGGTTCTG	GCAAATATTC
181	TGAAATGAGC	TGTTGACAAT	TAATCATCCG	GTCCGTATAA	TCTGTGGAAT	TGTGAGCGGG
241	ATAACAATTT	CATCGCGAGG	TACCAAGCTA	TCACAAGTTT	GTACAAAAAA	GCTGAACGAG
301	AAACGTAAAA	TGATATAAAT	ATCAATATAT	TAAATTAGAT	TTTGCATAAA	AAACAGACTA
361	CATAATACTG	TAAAAACAAA	CATATCCAGT	CACTATGGCG	GCCGCTAAGT	TGGCAGCATC
421	ACCCGACGCA	CTTTGCGCCG	AATAAATACC	TGTGACGGAA	GATCACTTCG	CAGAATAAAT
481	AAATCCTGGT	GTCCCTGTTG	ATACCGGGAA	GCCCTGGGCC	AACTTTTGGC	GAAAAATGAGA
541	CGTTGATCGG	CACGTAAGAG	GTTCCAACTT	TCACCATAAT	GAAATAAGAT	CACTACCGGG
601	CGTATTTTTT	GAGTTATCGA	GATTTTTCAGG	AGCTAAGGAA	GCTAAAAATG	AGAAAAAAT
661	CACTGGATAT	ACCACCGTTG	ATATATCCCA	ATGGCATCGT	AAAGAACATT	TTGAGGCATT
721	TCAGTCAGTT	GCTCAATGTA	CCTATAACCA	GACCGTTCAG	CTGGATATTA	CGGCCTTTTT
781	AAAGACCGTA	AAGAAAAATA	AGCACAGTTT	TTATCCGGCC	TTTATTACAA	TTCTTGCCCC
841	CCTGATGAAT	GCTCATCCGG	AATTCCTGAT	GGCAATGAAA	GACGGTGAGC	TGGTGATATG
901	GGATAGTGTT	CACCTTGTTT	ACACCGTTTT	CCATGAGCAA	ACTGAAACGT	TTTCATCGCT
961	CTGGAGTGAA	TACCACGACG	ATTTCCGGCA	GTTTCTACAC	ATATATTCGC	AAGATGTGGC
1021	GTGTTACGGT	GAAAACCTGG	CCTATTTCCT	TAAAGGGTTT	ATTGAGAATA	TGTTTTTCGT
1081	CTCAGCCAAT	CCCTGGGTGA	GTTTCACCAG	TTTTGATTTA	AACGTGGCCA	ATATGGACAA
1141	CTTCTTCGCC	CCCGTTTTCA	CCATGGGCAA	ATATTATACG	CAAGGCGACA	AGGTGCTGAT
1201	GCCGTGGCG	ATTCAGGTTC	ATCATGCCGT	CTGTGATGGC	TTCCATGTCT	GCAGAATGCT
1261	TAATGAATTA	CAACAGTACT	GCGATGAGTG	GCAGGGCGGG	GCGTAAACGC	GTGGATCCGG
1321	CTTACTAAAA	GCCAGATAAC	AGTATGCGTA	TTTGCGCGCT	GATTTTTGCG	GTATAAGAAT
1381	ATATACTGAT	ATGTATACCC	GAAGTATGTC	AAAAAGAGGT	GTGCTATGAA	GCAGCGTATT
1441	ACAGTGACAG	TTGACAGCGA	CAGCTATCAG	TTGCTCAAGG	CATATATGAT	GTCAATATCT
1501	CCGGTCTGGT	AAGCACAACC	ATGCAGAATG	AAGCCCGTCG	TCTGCGTGCC	GAACGCTGGA
1561	AAGCGGAAAA	TCAGGAAGGG	ATGGCTGAGG	TCGCCCCGTT	TATTGAAATG	AACGGCTCTT
1621	TTGCTGACGA	GAACAGGGAC	TGGTGAAATG	CAGTTTAAGG	TTTACACCTA	TAAAAGAGAG
1681	AGCCGTTATC	GTCTGTTTGT	GGATGTACAG	AGTGATATTA	TTGACACGCC	CGGGCGACGG
1741	ATGGTGATCC	CCCTGGCCAG	TGCACGTCTG	CTGTGAGATA	AAGTCTCCCG	TGAACTTTAC
1801	CCGGTGGTGC	ATATCGGGGA	TGAAAGCTGG	CGCATGATGA	CCACCGATAT	GGCCAGTGTG
1861	CCGGTCTCCG	TTATCGGGGA	AGAAGTGGCT	GATCTCAGCC	ACCGCGAAAA	TGACATCAAA
1921	AACGCCATTA	ACCTGATGTT	CTGGGGAATA	TAAATGTCAG	GCTCCCTTAT	ACACAGCCAG
1981	TCTGCAGGTC	GACCATAGTG	ACTGGATATG	TTGTGTTTTA	CAGTATTATG	TAGTCTGTTT
2041	TTTATGCAAA	ATCTAATTTA	ATATATTGAT	ATTTATATCA	TTTTACGTTT	CTCGTTTCAGC
2101	TTTCTTGTAC	AAAGTGGTGA	TAGCTTGGCT	GTTTTGGCGG	ATGAGAGAAG	ATTTTCAGCC
2161	TGATACAGAT	TAAATCAGAA	CGCAGAAGCG	GTCTGATAAA	ACAGAAATTTG	CCTGGCGGCA
2221	GTAGCGCGGT	GGTCCACCT	GACCCCATGC	CGAACTCAGA	AGTGAAACGC	CGTAGCGCCG
2281	ATGGTAGTGT	GGGGTCTCCC	CATGCGAGAG	TAGGGAACTG	CCAGGCATCA	AATAAAACGA
2341	AAGGCTCAGT	CGAAAGACTG	GGCCTTTCGT	TTTATCTGTT	GTTTGTCCGT	GAACGCTCTC
2401	CTGAGTAGGA	CAAAATCCGCC	GGGAGCGGAT	TTGAACGTTG	CGAAGCAACG	GCCCGGAGGG
2461	TGGCGGGCAG	GACGCCCCCG	ATAAACTGCC	AGGCATCAAA	TTAAGCAGAA	GGCCATCCTG
2521	ACGGATGGCC	TTTTTGCGTT	TCTACAAACT	CTTTTTGTTT	ATTTTCTTAA	ATACATTCAA-

FIGURE 21B

34/240

```

2581 ATATGTATCC GCTCATGAGA CAATAACCCCT GATAAATGCT TCAATAATAT TGAAAAAGGA
2641 AGAGTATGAG TATTCAACAT TTCCGTGTCT CCCTTATTCC CTTTTTTGCG GCATTTTGCC
2701 TTCCTGTTTT TGCTCACCCA GAAACGCTGG TGAAAGTAAA AGATGCTGAA GATCAGTTGG
2761 GTGCACGAGT GGGTTACATC GAACTGGATC TCAACAGCGG TAAGATCCTT GAGAGTTTTT
2821 GCCCCGAAGA ACGTTTTCCA ATGATGAGCA CTTTTAAAGT TCTGCTATGT GCGCGGTAT
2881 TATCCCCTGT TGACCCCGGG CAAGAGCAAC TCGGTGCGCG CATACTATAT TCTCAGAATG
2941 ACTTGGTTGA GTACTCACCA GTCACAGAAA AGCATCTTAC GGATGGCATG ACAGTAAGAG
3001 AATTATGCAG TGCTGCCATA ACCATGAGTG ATAACACTGC GGCCAACTTA CTTCTGACAA
3061 CGATCGGAGG ACCGAAGGAG CTAACCGCTT TTTTGACAAA CATGGGGGAT CATGTTAACTC
3121 GCCTTGATCG TTGGGAACCG GAGCTGAATG AAGCCATACC AAACGACGAG CGTGACACCA
3181 CGATGCCTAC AGCAATGGCA ACAACGTTGC GCAACTATT AACTGGCGAA CTACTTACTC
3241 TAGCTTCCCG GCAACAATTA ATAGACTGGA TGGAGGCGGA TAAAGTTGCA GGACCACTTC
3301 TGCGCTCGGC CCTTCCGGCT GGCTGGTTTA TTGCTGATAA ATCTGGAGCC GGTGAGCGTG
3361 GGTCTCGCGG TATCATTGCA GCACTGGGGC CAGATGGTAA GCCCTCCCGT ATCGTAGTTA
3421 TCTACACGAC GGGGAGTCAG GCAACTATGG ATGAACGAAA TAGACAGATC GCTGAGATAG
3481 GTGCCTCACT GATTAAAGCAT TGGTAACTGT CAGACCAAGT TTACTIONAT CATGTTTAA
3541 TTGATTTAAA ACTTCATTTT TAATTTAAAA GGATCTAGGT GAAGATCCTT TTTGATAATC
3601 TCATGACCAA AATCCCTTAA CGTGAGTTTT CGTTCCACTG AGCGTCAGAC CCCGTAGAAA
3661 AGATCAAAGG ATCTTCTTGA GATCCTTTTT TTCTGCGCGT AATCTGCTGC TTGCAAAACA
3721 AAAAACCACC GCTACCAGCG GTGGTTTTGT TGCCGGATCA AGAGCTACCA ACTCTTTTTT
3781 CGAAGGTAAC TGGCTTCAGC AGAGCGCAGA TACCAAATAC TGTCTTCTA GTGTAGCCGT
3841 AGTTAGGCCA CCACTTCAAG AACTCTGTAG CACCGCTAC ATACCTCGCT CTGCTAATCC
3901 GTTACCAGT GGCTGCTGCC AGTGGCGATA AGTCGTGTCT TACCGGGTTG GACTCAAGAC
3961 GATAGTTACC GGATAAGGCG CAGCGGTCGG GCTGAACGGG GGGTTCGTGC ACACAGCCCA
4021 GCTTGGAGCG AACGACCTAC ACCGAAGTGA GATACCTACA GCGTGAGCTA TGAGAAAGCG
4081 CCACGCTTCC CGAAGGGAGA AAGGCGGACA GGTATCCGGT AAGCGGCAGG GTCGGAACAG
4141 GAGAGCGCAC GAGGGAGCTT CCAGGGGGAA ACGCCTGGTA TCTTTATAGT CCTGTCCGGT
4201 TTCGCCACCT CTGACTTGAG CGTCGATTTT TGTGATGCTC GTCAGGGGGG CGGAGCCTAT
4261 GGAAAAACGC CAGCAACGCG GCCTTTTTAC GGTTCCTGGC CTTTTGCTGG CTTTTGCTC
4321 ACATGTTCTT TCCTGCGTTA TCCCCTGATT CTGTGGATAA CCGTATTACC GCCTTTGAGT
4381 GAGCTGATAC CGCTCGCCGC AGCCGAACGA CCGAGCGCAG CGAGTCAGTG AGCGAGGAAG
4441 CGGAAGAGCG CCTGATGCGG TATTTTCTCC TTACGATCT GTGCGGTATT TCACACCGCA
4501 TAATTTTGTT AAAATTCGCG TTAAATTTT GTTAAATCAG CTCATTTTTT AACCAATAGG
4561 CCGAAATCGG CAAAATCCCT TATAAATCAA AAGAATAGAC CGAGATAGGG TTGAGTGTG
4621 TTCCAGTTTG GAACAAGAGT CCACTATTAA AGAACGTGGA CTCCAACGTC AAAGGGCGAA
4681 AAACCGTCTA TCAGGCGCAT GGCCCACTAC GTGAACCATC ACCCTAATCA AGTTTTTTGG
4741 GGTGAGGTTG CCGTAAAGCA CTAATCGGA ACCCTAAAGG GAGCCCCCGA TTTAGAGCTT
4801 GACGGGGAAA GCCGGCGAAG GTGGCGAGAA AGGAAGGGAA GAAAGCGAAA GGAGCGGGCG
4861 CTAGGGCGCT GGCAAGTGTA GCGGTCACGC TGCGCGTAAC CACCACACCC GCCGCGCTTA
4921 ATGCGCCGCT ACAGGGCGCG TCCATTCCGC ATTCAAGGCTG CTATGGTGCA CTCTCAGTAC
4981 AATCTGCTCT GATGCCGCAT AGTTAAGCCA GTACCAGTCA CGTAGCGATA TCGGAGTGTA
5041 TACACTCCGC TATCGCTACG TGACTGGGTC ATGGCTGCGC CCCGACACCC GCCAACACCC
5101 GCTGACGCGC CCTGACGGGC TTGTCTGCTC CCGGCATCCG CTTACAGACA AGCTGTGACC
5161 GTCTCCGGGA GCTGCATGTG TCAGAGGTTT TCACCGTCAT CACCGAAACG CGCGAGGCAG
5221 CAGATCAATT CGCGCGCGAA GCGGAAGCGG CATGCATTTA CGTTGACACC ATCGAATGGT
5281 GCAAAACCTT TCGCGGTATG GCATGATAGC GCGCGGAAGA GAGTCAATTC AGGGTGGTGA
5341 ATGTGAAACC AGTAACGTTA TACGATGTCG CAGAGTATGC CGGTGTCTCT TATCAGACCG
5401 TTTCCGCGT GGTGAACGAG GCCAGCCAG TTTCTGCGAA AACCGGGGAA AAAGTGAAG
5461 TGGCGATGGC GGAGCTGAAT TACATTCCCA ACCGCGTGGC ACAACAACCTG CGGCGCAAAAC
5521 AGTCGTTGCT GATTGGCGTT GCCACCTCCA GTCTGGCCCT GCACGCGCCG TCGCAAATTG
5581 TCGCGGCGAT TAAATCTCGC GCCGATCAAC TGGGTGCCAG CGTGGTGGTG TCGATGGTAG
5641 AACGAAGCGG CGTCGAAGCC TGTAAGCGG CCGTGACAAA TCTTCTCGCG CAACCGCTCA
5701 GTGGGCTGAT CATTAACAT CCGCTGGATG ACCAGGATGC CATTGCTGTG GAAGCTGCCT
5761 GCATCAATGT TCCGGCGTTA TTTCTTGAT TCTCTGACCA GACACCCATC AACAGTATTA
5821 TTTTCTCCCA TGAAGACGGT ACGCGACTGG CCGTGAGCA TCTGGTCGCA TTGGGTCACC
5881 AGCAAATCGC GCTGTTAGCG GGCCCATTA GTTCTGTCTC GGCGCGCTG CGTCTGGCTG
5941 GCTGGCATAA ATATCTCACT CGCAATCAAA TTCAGCCGAT AGCGGAACGG GAAGGCGACT
6001 GGAGTGCCAT GTCCGGTTTT CAACAAACCA TGCAAATGCT GAATGAGGGC ATCGTTCCTCA-

```

Figure 21C

35/240

6061 CTGCGATGCT GGTGCCAAC GATCAGATGG CGCTGGGCGC AATGCGCGCC ATTACCGAGT
6121 CCGGGCTGCG CGTTGGTGCG GATATCTCGG TAGTGGGATA CGACGATACC GAAGACAGCT
6181 CATGTTATAT CCCGCCGTTA ACCACCATCA AACAGGATTT TCGCCTGCTG GGGCAAACCA
6241 GCGTGGACCG CTTGCTGCAA CTCTCTCAGG GCCAGGCGGT GAAGGGCAAT CAGCTGTTGC
6301 CCGTCTCACT GGTGAAAAGA AAAACCACCC TGGCACCCAA TACGCAAACC GCCTCTCCCC
6361 GCGCGTTGGC CGATTCATTA ATGCAGCTGG CACGACAGGT TTCCCGACTG GAAAGCGGGC
6421 AGTGAGCGCA ACGCAATTAA TGTGAGTTAG CGCGAATTGA TCTG

FIGURE 21D

36/240

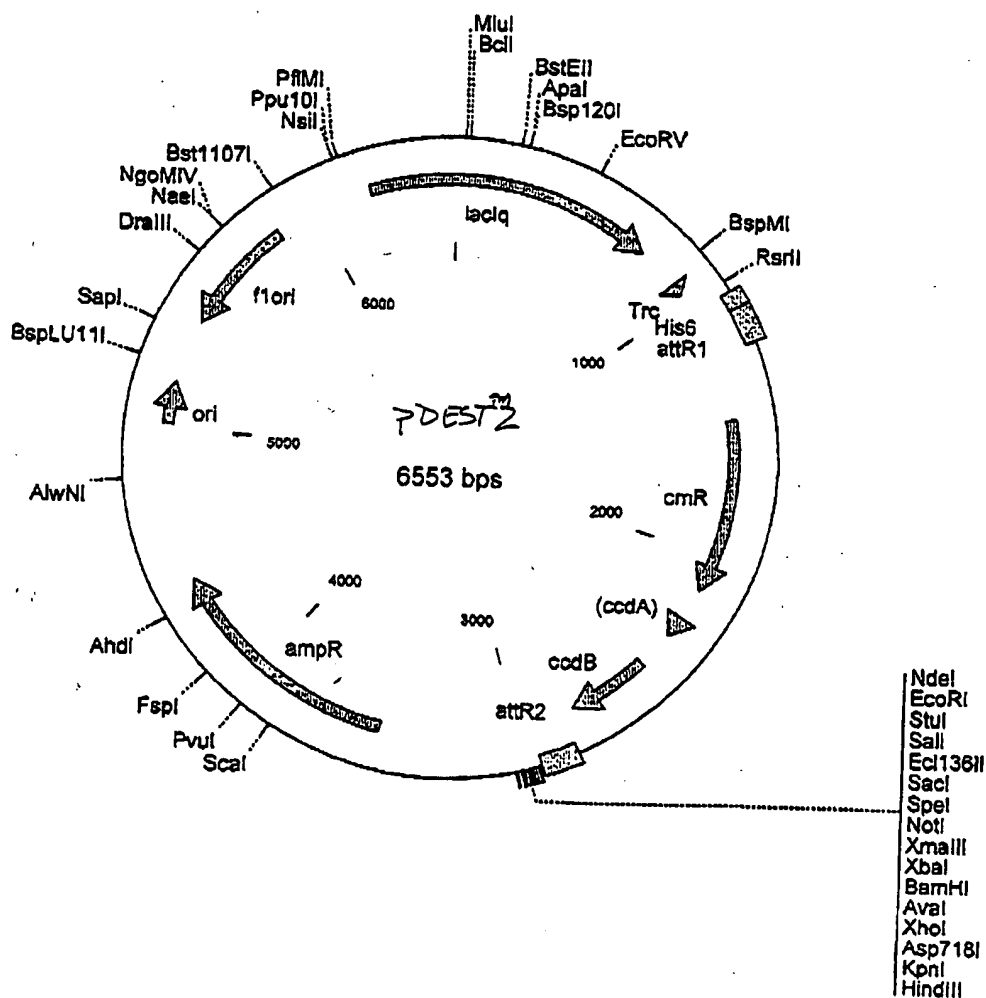
Figure 22A: pDEST2

His6 fusions in E. coli

970 aat att ctg aaa tga gct gtt gac aat tad tca tcc ggt ccg tat aat ctg
 tta taa gac ttt act cga caa ctg tta att agt agg cca ggc ata tta gac

1021 tgg aat tgt gag cgg ata aca att tca cac agg aaa cag acc atg ccg tac
 acc tta aca ctc gcc tat tgt taa agt gtg tcc ttt gtc tgg tac agc atg

1072 Tyr His His His His His His Glu Tle Trp Ser Tyr attR1
 tac cat cac cat cac cat cat ggt att aca agt tgg caa aga aaa gct gaa
 atg gta gtg gta gtg gta gtg ccg tag tgt tca aac atg ttt ctt caa ctt



37/240

pDEST2 6553 bp

<u>Location (Base Nos.)</u>		<u>Gene Encoded</u>
912..962		Trc
1223..1009		attR1
1473..2132		CmR
2252..2336		inactivated ccdA
2474..2779		ccdB
2820..2944		attR2
3509..4414		ampR
5015..5175		ori
5415..5852		flori (f1 intergenic region)
6225..752		lacIq

1	GGCGGTGCAC	AATCTTCTCG	CGCAACGCGT	CAGTGGGCTG	ATCATTAACT	ATCCGCTGGA
61	TGACCAGGAT	GCCATTGCTG	TGGAAGCTGC	CTGCACTAAT	GTTCCGGCGT	TATTTCTTGA
121	TGTCTCTGAC	CAGACACCCA	TCAACAGTAT	TATTTTCTCC	CATGAAGACG	GTACGCGACT
181	GGGCGTGGAG	CATCTGGTCG	CATTGGGTCA	CCAGCAAATC	GCGCTGTTAG	CGGGCCCAT
241	AAGTTCTGTC	TCGGCGCGTC	TGCGTCTGGC	TGGCTGGCAT	AAATATCTCA	CTCGCAATCA
301	AATTCAGCCG	ATAGCGGAAC	GGGAAGGCGA	CTGGAGTGCC	ATGTCCGGTT	TTCAACAAAC
361	CATGCAAATG	CTGAATGAGG	GCATCGTTCC	CACTGCGATG	CTGGTTGCCA	ACGATCAGAT
421	GGCGCTGGGC	GCAATGCGCG	CCATTACCGA	GTCCGGGCTG	CGCGTTGGTG	CGGATATCTC
481	GGTAGTGGGA	TACGACGATA	CCGAAGACAG	CTCATGTTAT	ATCCCGCCGT	CAACCACCAT
541	CAAACAGGAT	TTTCGCCTGC	TGGGGCAAAC	CAGCGTGGAC	CGCTTGCTGC	AACTCTCTCA
601	GGGCCAGGCG	GTGAAGGGCA	ATCAGCTGTT	GCCCGTCTCA	CTGGTGAAAA	GAAAAACCAC
661	CCTGGCACCC	AATACGCAA	CCGCCTCTCC	CCGCGCGTTG	GCCGATTCAT	TAATGCAGCT
721	GGCACGACAG	GTTTCCCGAC	TGGAAAGCGG	GCAGTGAGCG	CAACGCAATT	AATGTGAGTT
781	AGCGCGAATT	GATCTGGTTT	GACAGCTTAT	CATCGACTGC	ACGGTGCACC	AATGCTTCTG
841	GCGTCAGGCA	GCCATCGGAA	GCTGTGGTAT	GGCTGTGCAG	GTCGTAAATC	ACTGCATAAT
901	TCGTTGCGCT	CAAGGCGCAC	TCCCGTTCTG	GATAATGTTT	TTTGCGCCGA	CATCATAACG
961	GTTCTGGCAA	ATATTCTGAA	ATGAGCTGTT	GACAATTAAT	CATCCGGTCC	GTATAATCTG
1021	TGGAATTGTG	AGCGGATAAC	AATTTACAC	AGGAAACAGA	CCATGTCGTA	CTACCATCAC
1081	CATCACCATC	ACGGCATCAC	AAGTTTGATC	AAAAAAGCTG	AACGAGAAAC	GTAAATATGAT
1141	ATAAATATCA	ATATATTAAA	TTAGATTTTG	CATAAAAAAC	AGACTACATA	ATACTGTAAA
1201	ACACAACATA	TCCAGTCACT	ATGGCGGCGC	CTAAGTTGGC	AGCATCACCC	GACGCACTTT
1261	CGGCCGAATA	AATACCTGTG	ACGGAAGATC	ACTTCGCAGA	ATAAATAAAT	CCTGGTGCTC
1321	CTGTTGATAC	CGGGAAGCCC	TGGGCCAACT	TTTGCGGAAA	ATGAGACGTT	GATCGGCACG
1381	TAAGAGGTTT	CAACTTTCAC	CATAATGAAA	TAAGATCACT	ACCGGGCGTA	TTTTTTGAGT
1441	TATCGAGATT	TTCAGGAGCT	AAGGAAGCTA	AAATGGAGAA	AAAAATCACT	GGATATACCA
1501	CCGTTGATAT	ATCCCAATGG	CATCGTAAAG	AACATTTTGA	GGCATTTTCAG	TCAGTTGCTC
1561	AATGTACCTA	TAACCAGACC	GTTTCAGCTG	ATATTACGGC	CTTTTAAAG	ACCGTAAAGA
1621	AAAATAAGCA	CAAGTTTAT	CCGGCCTTTA	TTCACATTCT	TGCCCGCCTG	ATGAATGCTC
1681	ATCCGGAATT	CCGTATGGCA	ATGAAAGACG	GTGAGCTGGT	GATATGGGAT	AGTGTTACCC
1741	CTTGTTACAC	CGTTTTCCAT	GAGCAAACCTG	AAACGTTTTT	ATCGCTCTGG	AGTGAATACC
1801	ACGACGATTT	CCGGCAGTTT	CTACACATAT	ATTGCGCAAG	TGTGGCGTGT	TACGGTGAAA
1861	ACCTGGCCTA	TTTCCCTAAA	GGGTTTATTG	AGAATATGTT	TTTCGTCTCA	GCCAATCCCT
1921	GGGTGAGTTT	CACCAAGTTT	GATTTAAACG	TGGCCAATAT	GGACAACTTC	TTCCGCCCCG
1981	TTTTACCAT	GGGCAAATAT	TATACGCAAG	GCGACAAGGT	GCTGATGCCG	CTGGCGATT
2041	AGGTTTCATCA	TGCCGTCTGT	GATGGCTTCC	ATGTCGGCAG	AATGCTTAAT	GAATTACAAC
2101	AGTATGCGA	TGAGTGGCAG	GGCGGGCGGT	AAACGCGTGG	ATCCGGCTTA	CTAAAAGCCA
2161	GATAACAGTA	TGCGTATTGT	CGCGCTGATT	TTTGCGGTAT	AAGAATATAT	ACTGATATGT
2221	ATACCCGAAG	TATGTCAAAA	AGAGGTGTGC	TATGAAGCAG	CGTATTACAG	TGACAGTTGA
2281	CAGCGACAGC	TATCAGTTGC	TCAAGGCATA	TATGATGTCA	ATATCTCCGG	TCTGGTAAAG
2341	ACAACCATGC	AGAATGAAGC	CCGTGCTCTG	CGTGCCGAAC	GCTGGAAAGC	GGAAATTCAG
2401	GAAGGGATGG	CTGAGGTCTG	CCGGTTTATT	GAAATGAACG	GCTCTTTTGC	TGACGAGAAC
2461	AGGGACTGGT	GAAATGCAGT	TTAAGGTTTA	CACCTATAAA	AGAGAGAGCC	GTTATCGTCT
2521	GTTTGTGGAT	GTACAGAGTG	ATATTATTGA	CACGCCCGGG	CGACGGATGG	TGATCCCCCT-

FIGURE 22B

38/240

2581 GGCCAGTGCA CGTCTGCTGT CAGATAAAGT CTCCCGTGAA CTTTACCCGG TGGTGCATAT
2641 CGGGGATGAA AGCTGGCGCA TGATGACCAC CGATATGGCC AGTGTGCCCG TCTCCGTTAT
2701 CGGGGAAGAA GTGGCTGATC TCAGCCACCG CGAAAATGAC ATCAAAAACG CCATTAACCT
2761 GATGTTCTGG GGAATATAAA TGTCAGGCTC CTTTATACAC AGCCAGTCTG CAGGTCGACC
2821 ATAGTGAAGT GATATGTTGT GTTTTACAGT ATTATGTAGT CTGTTTTTTA TGCAAAATCT
2881 AATTTAATAT ATTGATATTT ATATCATTTT ACGTTTCTCG TTCAGCTTTC TTGTACAAAG
2941 TGGTGATGCC CATATGGGAA TTCAAAGGCC TACGTGACG AGCTCACTAG TCGCGGCCGC
3001 TTCTAGAGGA TCCCTCGAGG CATGCGGTAC CAAGCTTGCC TGTTTTGGCG GATGAGAGAA
3061 GATTTTCAGC CTGATACAGA TTAAATCAGA ACGCAGAAGC GGTCTGATAA AACAGAATTT
3121 GCCTGGCGGC AGTAGCGCGG TGGTCCACCC TGACCCCATG CCGAACTCAG AAGTGAAACG
3181 CCGTAGCGCC GATGTTAGTG TGGGGTCTCC CCATGCGAGA GTAGGGAACT GCCAGGCATC
3241 AAATAAAACG AAAGGCTCAG TCGAAAGACT GGGCCTTTCC TTTTATCTGT TGTGTGTCGG
3301 TGAACGCTCT CCTGAGTAGG ACAAATCCGC CGGGAGCGGA TTTGAACGTT GCGAAGCAAC
3361 GGCCCGGAGG GTGGCGGGCA GGACGCCCGC CATAAACTGC CAGGCATCAA ATTAAGCAGA
3421 AGGCCATCCT GACGGATGGC CTTTTTGCGT TTCTACAAAC TCTTTTGTGTT TAATTTTCTA
3481 AATACATTCA AATATGTATC CGCTCATGAG ACAATAACCC TGATAAATGC TTCAATAATA
3541 TTGAAAAAGG AAGAGTATGA GTATTCAACA TTTCCGTGTC GCCCTTATTC CCTTTTGTG
3601 GGCATTTTGC CTTCTGTTTT TTGCTCACCC AGAAACGCTG GTGAAAGTAA AAGATGCTGA
3661 AGATCAGTTG GGTGCACGAG TGGGTTACAT CGAACTGGAT CTCAACAGCG GTAAGATCCT
3721 TGAGAGTTTT CGCCCCGAAG AACGTTTTTC AATGATGAGC ACTTTTAAAG TTCTGCTATG
3781 TGGCGCGGTA TTATCCCGTG TTGACGCCGG GCAAGAGCAA CTCGGTCGCG GCATACACTA
3841 TTCTCAGAAAT GACTTGTTG AGTACTCACC AGTCACAGAA AAGCATCTTA CCGATGGCAT
3901 GACAGTAAGA GAATTATGCA GTGCTGCCAT AACCATGAGT GATAACACTG CGGCCAACTT
3961 ACTTCTGACA ACGATCGGAG GACCGAAGGA GCTAACCGCT TTTTGCACA ACATGGGGGA
4021 TCATGTAACCT CGCCTTGATC GTTGGGAACC GGAGCTGAAT GAAGCCATAC CAAACGACGA
4081 GCGTGACACC ACGATGCCTA CAGCAATGGC AACAACTTG CGAAACTAT TAACTGGCGA
4141 ACTACTTACT CTAGCTTCCC GGCAACAATT AATAGACTGG ATGGAGGCGG ATAAAGTTGC
4201 AGGACCACTT CTGCGCTCGG CCCTTCCGGC TGGCTGGTTT ATTGCTGATA AATCTGGAGC
4261 CCGTGAGCGT GGGTCTCGCG GTATCATTGC AGCACTGGGG CCAGATGGTA AGCCCTCCCG
4321 TATCGTAGTT ATCTACACGA CGGGGAGTCA GGCAACTATG GATGAACGAA ATAGACAGAT
4381 CGCTGAGATA GGTGCTCAC TGATTAAGCA TTGGTAACTG TCAGACCAAG TTTACTCATA
4441 TATACTTTAG ATTGATTTAA AACTTCATTT TTAATTTAAA AGGATCTAGG TGAAGATCCT
4501 TTTTGATAAT CTCATGACCA AAATCCCTTA ACGTGAGTTT TCGTTCCACT GAGCGTCAGA
4561 CCCCGTAGAA AAGATCAAAG GATCTTCTTG AGATCCTTTT TTTCTGCGCG TAATCTGCTG
4621 CTTGCAACA AAAAACCAC CGCTACCAGC GGTGGTTTGT TTGCCGGATC AAGAGCTACC
4681 AACTCTTTTT CCGAAGGTAA CTGGCTTCAG CAGAGCGCAG ATACCAAATA CTGTCCTTCT
4741 AGTGTAGCCG TAGTTAGGCC ACCACTTCAA GAACTCTGTA GCACCGCCTA CATACTCGC
4801 TCTGCTAATC CTGTTACCAG TGGCTGCTGC CAGTGGCGAT AAGTCGTGTC TTACCGGGTT
4861 GGAATCAAGA CGATAGTTAC CGGATAAGGC GCAGCGGTG GGTGAACGG GGGGTTCTGT
4921 CACACAGCCC AGCTTGGAGC GAACGACCTA CACCGAACTG AGATACCTAC AGCGTGAGCT
4981 ATGAGAAAAG GCCACGCTTC CCGAAGGGAG AAAGGCGGAC AGGTATCCGG TAAGCGGCAG
5041 GGTGGAACA GGAGAGCGCA CGAGGGAGCT TCCAGGGGGA AACGCCTGGT ATCTTTATAG
5101 TCCTGTGCGG TTTGCCCACC TCTGACTTGA GCGTCGATTT TTGTGATGCT CGTCAGGGGG
5161 GCGGAGCCTA TGGAAAAACG CCAGCAACGC GGCTTTTTTA CGGTTCCCTG CCTTTTGCTG
5221 GCCTTTTGCT CACATGTTCT TTCCTGCGTT ATCCCTGAT TCTGTGGATA ACCGTATTAC
5281 CGCCTTTGAG TGAGCTGATA CCGCTCGCG CAGCCGAACG ACCGAGCGCA GCGAGTCAGT
5341 GAGCGAGGAA GCGGAAGAGC GCCTGATGCG GTATTTTCTC CTTACGCATC TGTGCGGTAT
5401 TTCACACCGC ATAATTTTGT TAAATTCGCG GTTAAATTTT TGTTAAATCA GCTCATTTTT
5461 TAACCAATAG GCCGAAATCG GCAAAATCCC TTATAAATCA AAAGAATAGA CCGAGATAGG
5521 GTTGAGTGT GTTCCAGTTT GGAACAAGAG TCCACTATTA AAGAACGTGG ACTCCAACGT
5581 CAAAGGGCGA AAAACCGTCT ATCAGGGCGA TGGCCCACTA CGTGAACCAT CACCTAATC
5641 AAGTTTTTTG GGGTCGAGGT GCCGTAAGC ACTAAATCGG AACCTAAAG GGAGCCCCCG
5701 ATTTAGAGCT TGACGGGGAA AGCCGGCGAA CGTGGCGAGA AAGGAAGGGA AGAAAGCGAA
5761 AGGAGCGGGC GCTAGGGCGC TGGCAAGTGT AGCGGTACG CTGCGCGTAA CCACCAACCC
5821 CGCCGCGCTT AATGCGCCGC TACAGGGCGC GTCCCATTCG CCATTACAGC TGCTATGGTG
5881 CACTCTCAGT ACAATCTGCT CTGATGCCGC ATAGTTAAGC CAGTATACAC TCCGCTATCG
5941 CTACGTGACT GGGTCATGGC TGCGCCCCGA CACCCGCCAA CACCCGCTGA CGCGCCTGA
6001 CGGGCTTGTC TGCTCCCGGC ATCCGCTTAC AGACAAGCTG TGACCGTCTC CGGGAGCTGC-

FIGURE 22C

6061 ATGTGTCAGA GGTTCACCC GTCATCACCG AAACGCGCGA GGCAGCAGAT CAATTCGCGC
6121 GCGAAGGCGA AGCGGCATGC ATTTACGTTG ACACCATCGA ATGGTGCAAA ACCTTTCGCG
6181 GTATGGCATG ATAGCGCCCG GAAGAGAGTC AATTCAGGGT GGTGAATGTG AAACCAGTAA
6241 CGTTATACGA TGTCGCAGAG TATGCCGGTG TCTCTTATCA GACCGTTTCC CGCGTGGTGA
6301 ACCAGGCCAG CCACGTTTCT GCGAAAACGC GGGAAAAAGT GGAAGCGGCG ATGGCGGAGC
6361 TGAATTACAT TCCCAACCGC GTGGCACAAC AACTGGCGGG CAAACAGTCG TTGCTGATTG
6421 GCGTTGCCAC CTCCAGTCTG GCCCTGCACG CGCCGTCGCA AATTGTCGCG GCGATTAAAT
6481 CTCGCGCCGA TCAACTGGGT GCCAGCGTGG TGGTGTGAT GGTAGAACGA AGCGGCGTCG
6541 AAGCCTGTAA AGC

40/240

Figure 23A: pDEST3

GST fusions in E. coli

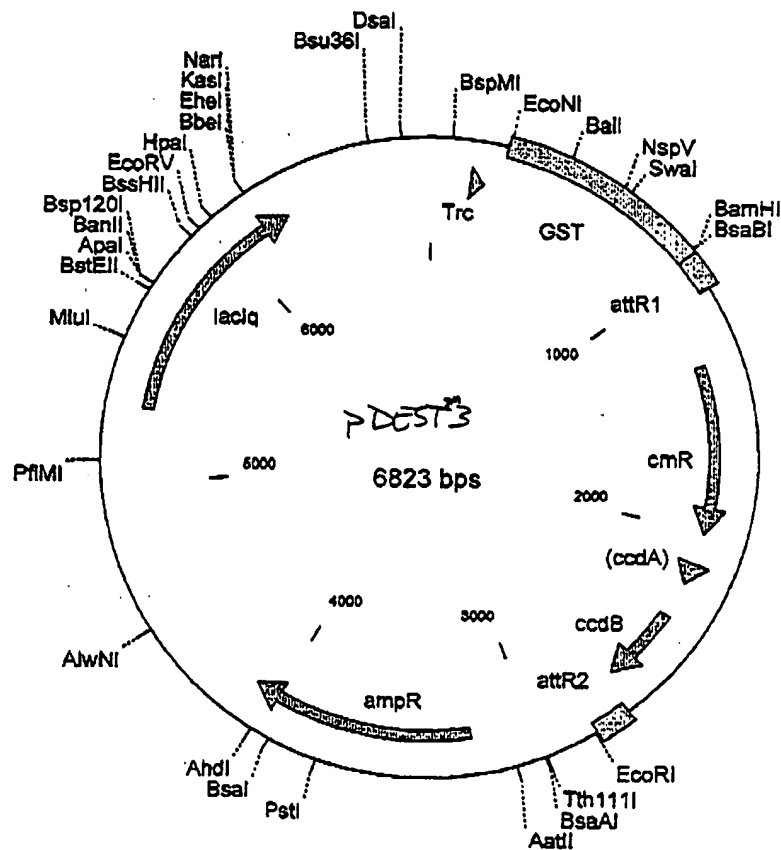
154 cgg ttc tgg caa ata ttc tga aat gag ctg ⁻³⁵ ctg aca att aat cat cgg ctc
 gcc aag acc gtt tat aag act tta ctc gac aac tgt taa tta gta gcc gag

205 ⁻¹⁰ gca taa ^{→ mRNA} ggt gtg gaa ttg tga gcg gat aac aat ttc aca cag gaa aca gta
 cat att aca cac ctt aac act cgc cta ttg tta aag tgt gtc ctt tgt cat

256 ^{M S P I L} ^{→ GST} ttc atg tcc cct ata cta ggt tat tgg aaa att aag ggc ctt gtg caa ccc
 aag tac agg gga tat gat cca ata acc ttt taa ttc ccg gaa cac gtt ggg

919 ^{" GST} [→] ctg gtt ccg cgt gga tct cgt cgt gca tct gtt gga tcc cca tca ^{T S} aca agt
 gac caa ggc gca cct aga gca gca cgt aga caa cct agg ggt agt tgt tca

970 ^{H Y K K} ~~ctg tac aac aca gct gaa cga gaa acg taa aat gat ata aat aat aat ata~~
~~aac atg ttt ttc cga cct gct cct tgc att tta cta tat tta tag tta tat~~



41/240

pDEST3 6823 bp

<u>Location (Base Nos.)</u>		<u>Gene Encoded</u>
150..200		Trc
1087..963		attR1
1337..1996		CmR
2116..2200		inactivated ccdA
2338..2643		ccdB
2684..2808		attR2
3231..4091		ampR
5295..6254		lacIq

1	ACGTTATCGA	CTGCACGGTG	CACCAATGCT	TCTGGCGTCA	GGCAGCCATC	GGAAGCTGTG
61	GTATGGCTGT	GCAGGTCGTA	AATCACTGCA	TAATTCGTGT	CGCTCAAGGC	GCACTCCCGT
121	TCTGGATAAT	GTTTTTTGCG	CCGACATCAT	AACGGTCTCTG	GCAAATATTC	TGAAATGAGC
181	TGTTGACAAAT	TAATCATCGG	CTCGTATAAT	GTGTGGAATT	GTGAGCGGAT	AACAATTTC
241	CACAGGAAAC	AGTATTCATG	TCCCTATAC	TAGGTTATTG	GAAAATTAAG	GGCCTTGTGC
301	AACCCACTCG	ACTTCTTTTG	GAATATCTTG	AAGAAAAATA	TGAAGAGCAT	TTGTATGAGC
361	GCGATGAAGG	TGATAAATGG	CGAAACAAAA	AGTTTGAATT	GGGTTTGGAG	TTTCCCAATC
421	TTCTTTATTA	TATTGATGGT	GATGTTAAAT	TAACACAGTC	TATGGCCATC	ATACGTTATA
481	TAGCTGACAA	GCACAACATG	TTGGGTGGTT	GTCCAAAAGA	GCGTGCAGAG	ATTTCAATGC
541	TTGAAGGAGC	GGTTTTGGAT	ATTAGATACG	GTGTTTCGAG	AATTGCATAT	AGTAAAGACT
601	TTGAAACTCT	CAAAGTTGAT	TTTCTTAGCA	AGCTACCTGA	AATGCTGAAA	ATGTTTGAAG
661	ATCGTTTATG	TCATAAAACA	TATTTAAATG	GTGATCATGT	AACCCATCCT	GACTTCATGT
721	TGTATGACGC	TCTTGATGTT	GTTTTATACA	TGGACCCAAT	GTGCCTGGAT	GCGTTCCCAA
781	AATTAGTTTG	TTTTAAAAAA	CGTATTGAAG	CTATCCACAC	AATTGATAAG	TACTTGAAAT
841	CCAGCAAGTA	TATAGCATGG	CCTTTCGAGG	GCTGGCAAGC	CACGTTTGGT	GGTGGCGACC
901	ATCCTCCAAA	ATCGGATCTG	GTTCCGCGTG	GATCTCGTCG	TGCATCTGTT	GGATCCCCAT
961	CAACAAGTTT	GTACAAAAAA	GCTGAACGAG	AAACGTAAAA	TGATATAAAT	ATCAATATAT
1021	TAAATTAGAT	TTTGCAATAA	AAACAGACTA	CATAATACTG	TAAAACACAA	CATATCCAGT
1081	CACTATGGCG	GCCGCTAAGT	TGGCAGCATC	ACCCGACGCA	CTTTGCGCCG	AATAAATACC
1141	TGTGACGGAA	GATCACTTCG	CAGAATAAAT	AAATCCTGGT	GTCCCTGTTG	ATACCGGGAA
1201	GCCCTGGGCC	AACTTTGGGC	GAAAATGAGA	CGTTGATCGG	CACGTAAGAG	GTTCCAACTT
1261	TCACCATAAT	GAAATAAGAT	CACTACCGGG	CGTATTTTTT	GAGTTATCGA	GATTTTCAGG
1321	AGCTAAGGAA	GCTAAAATGG	AGAAAAAAT	CACTGGATAT	ACCACCGTTG	ATATATCCCA
1381	ATGGCATCGT	AAAGAACATT	TTGAGGCATT	TCAGTCAGTT	GCTCAATGTA	CCTATAACCA
1441	GACCGTTCAG	CTGGATATTA	CGGCCTTTTT	AAAGACCGTA	AAGAAAAATA	AGCACAGGTT
1501	TTATCCGGCC	TTTATTCACA	TTCTTGCCCC	CCTGATGAAT	GCTCATCCGG	AATTCGGTAT
1561	GGCAATGAAA	GACGGTGAGC	TGGTGATATG	GGATAGTGTT	CACCCTTGTT	ACACCGTTTT
1621	CCATGAGCAA	ACTGAAACGT	TTTCATCGCT	CTGGAGTGAA	TACCACGACG	ATTTCCGGCA
1681	GTTTCTACAC	ATATATTCGC	AAGATGTGGC	GTGTTACGGT	GAAAACCTGG	CCTATTTCCC
1741	TAAAGGGTTT	ATTGAGAATA	TGTTTTTCGT	CTCAGCCAAT	CCCTGGGTGA	GTTTCACCAG
1801	TTTTGATTTA	AACGTGGCCA	ATATGGACAA	CTTCTTCGCC	CCCGTTTTCA	CCATGGGCAA
1861	ATATTATACG	CAAGGCGACA	AGGTGCTGAT	GCCGCTGGCG	ATTCAGGTTT	ATCATGCCGT
1921	CTGTGATGGC	TTCCATGTCG	GCAGAATGCT	TAATGAATTA	CAACAGTACT	GCGATGAGTG
1981	GCAGGGCGGG	GCGTAAAGAT	CTGGATCCGG	CTTACTAAAA	GCCAGATAAC	AGTATGCGTA
2041	TTTGCGCGCT	GATTTTTGCG	GTATAAGAAT	ATATACTGAT	ATGTATACCC	GAAGTATGTC
2101	AAAAAGAGGT	GTGCTATGAA	GCAGCGTATT	ACAGTGACAG	TTGACAGCGA	CAGCTATCAG
2161	TGTCTCAAGG	CATATATGAT	GCAATATCT	CCGGTCTGGT	AAGCACAACC	ATGCAGAATG
2221	AAGCCCGTCG	TCTGCGTGCC	GAACGCTGGA	AAGCGGAAAA	TCAGGAAGGG	ATGGCTGAGG
2281	TCGCCCCGTT	TATTGAAATG	AACGGCTCTT	TTGCTGACGA	GAACAGGGAC	TGGTGAAATG
2341	CAGTTTAAAG	TTTACACCTA	TAAAAGAGAG	AGCCGTTATC	GTCTGTTTGT	GGATGTACAG
2401	AGTGATATTA	TTGACACGCC	CGGGCGACGG	ATGGTGATCC	CCCTGGCCAG	TGCACGCTCG
2461	CTGTCAGATA	AAGTCTCCCG	TGAACTTTAC	CCGGTGGTGC	ATATCGGGGA	TGAAAGCTGG
2521	CCCATGATGA	CCACCGATAT	GGCCAGTGTC	CCGGTCTCCG	TTATCGGGGA	AGAAGTGGCT
2581	GATCTCAGCC	ACCGCGAAAA	TGACATCAAA	AACGCCATTA	ACCTGATGTT	CTGGGGGAATA
2641	TAAATGTCAG	GCTCCCTTAT	ACACAGCCAG	TCTGCAGGTC	GACCATAGTG	ACTGGATATG-

FIGURE 23B

42/240

2701 TTGTGTTTTA CAGTATTATG TAGTCTGTTT TTTATGCAAA ATCTAATTTA ATATATTGAT
 2761 ATTTATATCA TTTTACGTTT CTCGTTTCAGC TTTCTTGATC AAAGTGGTTG ATGGGAATTC
 2821 ATCGTGACTG ACTGACGATC TGCCTCGCGC GTTTCGGTGA TGACGGTGAA AACCTCTGAC
 2881 ACATGCAGCT CCCGGAGACG GTCACAGCTT GTCTGTAAGC GGATGCCGGG AGCAGACAAAG
 2941 CCCGTCAGGG CGCGTCAGCG GGTGTTGGCG GGTGTCGGGG CGCAGCCATG ACCCAGTCAC
 3001 GTAGCGATAG CGGAGTGTAT AATTCCTGAA GACGAAAGGG CCTCGTGATA CGCCTATTTT
 3061 TATAGGTTAA TGTCATGATA ATAATGGTTT CTTAGACGTC AGGTGGCACT TTTCCGGGAA
 3121 ATGTGCGCGG AACCCCTATT TGTTTATTTT TCTAAATACA TTCAAATATG TATCCGCTCA
 3181 TGAGACAATA ACCCTGATAA ATGCTTCAAT AATATTGAAA AAGGAAGAGT ATGAGTATTC
 3241 AACATTTCCG TGTCGCCCTT ATTCCTTTT TTGCGGCATT TTGCCTTCCT GTTTTTGCTC
 3301 ACCCAGAAAC GCTGGTGAAA GTAAAAGATG CTGAAGATCA GTTGGGTGCA CGAGTGGGTT
 3361 ACATCGAAGT GGATCTCAAC AGCGGTAAGA TCCTTGAGAG TTTTCGCCCC GAAGAACGTT
 3421 TTCCAATGAT GAGCACTTTT AAAGTTCTGC TATGTGGCGC GGTATTATCC CGTGTGACG
 3481 CCGGGCAAGA GCAACTCGGT CGCCGCATAC ACTATTCTCA GAATGACTTG GTTGAGTACT
 3541 CACCAGTCAC AGAAAAGCAT CTTACGGATG GCATGACAGT AAGAGAATTA TGCAGTGCTG
 3601 CCATAACCAT GAGTGATAAC ACTCGGGCCA ACTTACTTCT GACAACGATC GGAGGACCGA
 3661 AGGAGCTAAC CGCTTTTTTG CACAACATGG GGGATCATGT AACTCGCCTT GATCGTTGGG
 3721 AACCGGAGCT GAATGAAGCC ATACCAAACG ACGAGCGTGA CACCACGATG CCTGCAGCAA
 3781 TGCAACAAC GTTGCAGCAA CTATTAAGTG GCGAAGTACT TACTCTAGCT TCCCGGCAAC
 3841 AATTAATAGA CTGGATGGAG GCGGATAAAG TTGCAGGACC ACTTCTGCGC TCGGCCCTTC
 3901 CGGCTGGCTG GTTTATTGCT GATAAATCTG GAGCCGGTGA GCGTGGGTCT CGCGGTATCA
 3961 TTGCAGCACT GGGGCCAGAT GGTAAAGCCT CCCGTATCGT AGTTATCTAC ACGACGGGGA
 4021 GTCAGGCAAC TATGGATGAA CGAAATAGAC AGATCGCTGA GATAGGTGCC TCACGTATTA
 4081 AGCATTGGTA ACTGTCAGAC CAAGTTTACT CATATATACT TTAGATTGAT TTAAACTTTC
 4141 ATTTTAAATT TAAAAGGATC TAGGTGAAGA TCCTTTTGA TAATCTCATG ACCAAAATCC
 4201 CTTAACGTGA GTTTTCGTTT CACTGAGCGT CAGACCCCGT AGAAAAGATC AAAGGATCTT
 4261 CTTGAGATCC TTTTTTCTG CGCGTAATCT GCTGCTTGCA AACAAAAAAA CCACCGCTAC
 4321 CAGCGGTGGT TTGTTTGCCG GATCAAGAGC TACCAACTCT TTTTCCGAAG GTAACGGCT
 4381 TCAGCAGAGC GCAGATACCA AATACTGTCC TTCTAGTGTA GCCGTAGTTA GGCCACCACT
 4441 TCAAGAACTC TGTAGCACCG CCTACATACC TCGCTCTGCT AATCCTGTTA CCAGTGGCTG
 4501 CTGCCAGTGG CGATAAGTCG TGTCTTACCG GGTGGACTC AAGACGATAG TTACCGGATA
 4561 AGGCGCAGCG GTCGGGCTGA ACGGGGGGTT CGTGACACACA GCCCAGCTTG GAGCGAACGA
 4621 CCTACACCGA ACTGAGATAC CTACAGCGTG AGCTATGAGA AAGCGCCACG CTTCCCGAAG
 4681 GGAGAAAGGC GGACAGGTAT CCGGTAAGCG GCAGGGTCGG AACAGGAGAG CGCACGAGGG
 4741 AGCTTCCAGG GGGAAACGCC TGGTATCTTT ATAGTCCTGT CGGGTTTCGC CACCTCTGAC
 4801 TTGAGCGTCG ATTTTGTGA TGCTCGTCAG GGGGGCGGAG CCTATGGAAA AACGCCAGCA
 4861 ACGCGGCCCT TTTACGGTTC CTGGCCTTTT GCTGGCCTTT TGCTCACATG TTCTTCTCTG
 4921 CGTTATCCCC TGATTCTGTG GATAACCGTA TTACCGCCTT TGAGTGAGCT GATACCGCTC
 4981 GCCGCAGCCG AACGACCGAG CGCAGCGAGT CAGTGAGCGA GGAAGCGGAA GAGCGCCTGA
 5041 TGCGGTATTT TCTCCTTACG CATCTGTGCG GTATTTTACA CCGCATAAAT TCCGACACCA
 5101 TCGAATGGTG CAAAACCTTT CGCGGTATGG CATGATAGCG CCCGGAAGAG AGTCAATTCA
 5161 GGGTGGTGAA TGTGAAACCA GTAACGTTAT ACGATGTCGC AGAGTATGCC GGTGTCTCTT
 5221 ATCAGACCGT TTCCCGCGTG GTGAACCAGG CCAGCCACGT TTCTGCGAAA ACGCGGGAAA
 5281 AAGTGGAAGC GGCGATGGCG GAGCTGAATT ACATTCCCAA CCGCGTGGCA CAACAACTGG
 5341 CGGGCAAACA GTCGTTGCTG ATTGGCGTTG CCACCTCCAG TCTGGCCCTG CACGCGCCGT
 5401 CGCAAATTGT CGCGGCGATT AAATCTCGCG CCGATCAACT GGGTGCCAGC GTGGTGGTGT
 5461 CGATGGTAGA ACGAAGCGGC GTCGAAGCCT GTAAAGCGGC GGTGCACAAAT CTTCTCGCGC
 5521 AACGCGTCAG TGGGCTGATC ATTAACATATC CGCTGGATGA CCAGGATGCC ATTGCTGTGG
 5581 AAGCTGCCTG CACTAATGTT CCGGCGTTAT TTCTTGATGT CTCTGACCAG ACACCGTATG
 5641 ACAGTATTAT TTTCTCCCAT GAAGACGGTA CGCGACTGGG CGTGAGCAT CTGGTCGCAT
 5701 TGGGTACCA GCAAATCGCG CTGTTAGCGG GCCCATTAAG TTCTGTCTCG GCGCGTCTGC
 5761 GTCTGGCTGG CTGGCATAAA TATCTCACTC GCAATCAAAT TCAGCCGATA GCGGAACGGG
 5821 AAGGCGACTG GAGTGCCATG TCCGGTTTTC AACAAACCAT GCAAATGCTG AATGAGGGCA
 5881 TCGTTCCAC TGCGATGCTG GTTGCCAACG ATCAGATGGC GCTGGGCGCA ATGCGCGCCA
 5941 TTACCGAGTC CGGGCTGCGC GTTGGTGCGG ATATCTCGGT AGTGGGATAC GACGATACCG
 6001 AAGACAGCTC ATGTTATATC CCGCGTTAA CCACCATCAA ACAGGATTTT CGCCTGCTGG
 6061 GGCAAACAG CGTGGACCGC TTGCTGCAAC TCTCTCAGGG CCAGGCGGTG AAGGGCAATC
 6121 AGCTGTTGCC CGTCTCACTG GTGAAAAGAA AAACCACCCT GCGGCCAAT ACGCAAACCG-

FIGURE 23C

43/240

6181 CCTCTCCCCG CGCGTTGGCC GATTCATTAA TGCAGCTGGC ACGACAGGTT TCCCGACTGG
6241 AAAGCGGGCA GTGAGCGCAA CGCAATTAAT GTGAGTTAGC TCACTCATTG GGCACCCAG
6301 GCTTTACACT TTATGCTTCC GGCTCGTATG TTGTGTGGAA TTGTGAGCGG ATAACAATTT
6361 CACACAGGAA ACAGCTATGA CCATGATTAC GGATTCACTG GCCGTCGTTT TACAACGTCG
6421 TGA CTGGGAA AACCCTGGCG TTACCCA ACT TAATCGCCTT GCAGCACATC CCCCTTTTCGC
6481 CAGCTGGCGT AATAGCGAAG AGGCCCGCAC CGATCGCCCT TCCCAACAGT TGCGCAGCCT
6541 GAATGGCGAA TGGCGCTTTG CCTGGTTTCC GGCACCAGAA GCGGTGCCGG AAAGCTGGCT
6601 GGAGTGCGAT CTTCTGAGG CCGATACTGT CGTCGTCCCC TCAAAC TGGC AGATGCACGG
6661 TTACGATGCG CCCATCTACA CCAACGTAAC CTATCCCATT ACGGTCAATC CGCCGTTTGT
6721 TCCACGGAG AATCCGACGG GTTGTTACTC GCTCACATTT AATGTTGATG AAAGCTGGCT
6781 ACAGGAAGGC CAGACCGGAA TTATTTTGA TGGCGTTGGA ATT

FIGURE 23D

45/240

pDEST4 6964 bp

<u>Location (Base Nos.)</u>		<u>Gene Encoded</u>
964..1003		Trc
1577..1453		attR1
1827..2486		CmR
2606..2690		inactivated ccdA
2828..3133		ccdB
3174..3298		attR2
3872..4777		ampR
5378..5538		ori
5778..6215		flori (f1 intergenic region)
6587..704		lacIq

1	CTATCCGCTG	GATGACCAGG	ATGCCATTGC	TGTGGAAGCT	GCCTGCACTA	ATGTTCCGGC
61	GTTATTCTTT	GATGTCTCTG	ACCAGACACC	CATCAACAGT	ATTATTTTCT	CCCATGAAGA
121	CGGTACGCGA	CTGGGCGTGG	AGCATCTGGT	CGCATTGGGT	CACCAGCAAA	TCGCGCTGTT
181	AGCGGGCCCA	TTAAGTTCTG	TCTCGGCGCG	TCTGCGTCTG	GCTGGCTGGC	ATAAATATCT
241	CACTCGCAAT	CAAATTCAGC	CGATAGCGGA	ACGGGAAGGC	GACTGGAGTG	CCATGTCCGG
301	TTTTCAACAA	ACCATGCAAA	TGCTGAATGA	GGGCATCGTT	CCCACTGCGA	TGCTGGTTGC
361	CAACGATCAG	ATGGCGCTGG	GCGCAATGCG	CGCCATTACC	GAGTCCGGGC	TGCGCGTTGG
421	TGCGGATATC	TCGGTAGTGG	GATACGACGA	TACCGAAGAC	AGCTCATGTT	ATATCCCGCC
481	GTCAACCACC	ATCAAACAGG	ATTTTCGCCT	GCTGGGGCAA	ACCAGCGTGG	ACCGCTTGCT
541	GCAACTCTCT	CAGGGCCAGG	CGGTGAAGGG	CAATCAGCTG	TTGCCCGTCT	CACTGGTGAA
601	AAGAAAAACC	ACCCTGGCAC	CCAATACGCA	AACCGCCTCT	CCCCGCGCGT	TGGCCGATTTC
661	ATTAATGCAG	CTGGCAGCAC	AGGTTTCCCG	ACTGGAAAGC	GGGCAGTGAG	CGCAACGCAA
721	TTAATGTGAG	TTAGCGCGAA	TTGATCTGGT	TTGACAGCTT	ATCATCGACT	GCACGGTGCA
781	CCAATGCTTC	TGGCGTCAGG	CAGCCATCGG	AAGCTGTGGT	ATGGCTGTGC	AGGTCGTAAA
841	TCACTGCATA	ATTCTGTCTG	CTCAAGGCGC	ACTCCCGTTC	TGGATAATGT	TTTTTGCGCC
901	GACATCATAA	CGGTTCTGGC	AAATATTCTG	AAATGAGCTG	TTGACAATTA	ATCATCCGGT
961	CCGTATAATC	TGTGGAATTG	TGAGCGGATA	ACAATTTTAC	ACAGGAAACA	GACCATGGGT
1021	CATCATCATC	ATCATCACGA	TTACGATATC	CCAACGACCG	AAAACCTGTA	TTTTCAGGGC
1081	GCCCATATGA	GCGATAAAAT	TATTCACCTG	ACTGACGACA	GTTTTGACAC	GGATGTACTC
1141	AAAGCGGACG	GGGCGATCCT	CGTCGATTTC	TGGGCAGAGT	GGTGCGGTCC	GTGCAAAATG
1201	ATCGCCCCGA	TTCTGGATGA	AATCGCTGAC	GAATATCAGG	GCAAACTGAC	CGTTGCAAAA
1261	CTGAACATCG	ATCAAAACCC	TGGCACTGCG	CCGAAATATG	GCATCCGTGG	TATCCCGACT
1321	CTGCTGCTGT	TCAAAAACGG	TGAAGTGGCG	GCAACCAAAG	TGGGTGCACT	GTCTAAAGGT
1381	CAGTTGAAAG	AGTTCCTCGA	CGCTAACCTG	GCCGGTTCTG	GTTCTGGTGA	TGACGATGAC
1441	AAGGTACCCA	TCACAAGTTT	GTACAAAAAA	GCTGAACGAG	AAACGTAAAA	TGATATAAAT
1501	ATCAATATAT	TAAATTAGAT	TTTGCAATAA	AAACAGACTA	CATAATACTG	TAAAAACAA
1561	CATATCCAGT	CACTATGGCG	GCGCTAAGT	TGGCAGCATC	ACCCGACGCA	CTTTGCGCCG
1621	AATAAATACC	TGTGACGGAA	GATCACTTCG	CAGAATAAAT	AAATCCTGGT	GTCCCTGTTG
1681	ATACCGGGAA	GCCCTGGGCC	AACTTTTGGC	GAAAATGAGA	CGTTGATCGG	CACGTAAGAG
1741	GTTCCAACCT	TCACCATAAT	GAAATAAGAT	CACTACCGGG	CGTATTTTTT	GAGTTATCGA
1801	GATTTTTCAG	AGCTAAGGAA	GCTAAAATGG	AGAAAAAAAT	CACTGGATAT	ACCACCGTTG
1861	ATATATCCCA	ATGGCATCGT	AAAGAACATT	TTGAGGCATT	TCAGTCAGTT	GCTCAATGTA
1921	CCTATAACCA	GACCGTTCAG	CTGGATATTA	CGGCCTTTTT	AAAGACCGTA	AAGAAAAATA
1981	AGCACAAGTT	TTATCCGGCC	TTTATTCACA	TTCTTGCCCC	CCTGATGAAT	GCTCATCCGG
2041	AATTCCGTAT	GGCAATGAAA	GACGGTGAGC	TGGTGATATG	GGATAGTGTT	CACCCTTGTT
2101	ACACCGTTTT	CCATGAGCAA	ACTGAAACGT	TTTCATCGCT	CTGGAGTGAA	TACCACGACG
2161	ATTTCCGGCA	GTTTCTACAC	ATATATTTCG	AAGATGTGGC	GTGTTACGGT	GAAAACCTGG
2221	CCTATTTCCC	TAAAGGGTTT	ATTGAGAAAT	TGTTTTTCGT	CTCAGCCAAT	CCCTGGGTGA
2281	GTTTCACCAG	TTTTGATTTA	AACGTGGCCA	ATATGGACAA	CTTCTTCGCC	CCCGTTTTCA
2341	CCATGGGCAA	ATATTATACG	CAAGGCGACA	AGGTGCTGAT	GCCGCTGGCG	ATTACAGGTT
2401	ATCATGCCGT	CTGTGATGGC	TTCATGTTCG	GCAGAATGCT	TAATGAATTA	CAACAGTACT
2461	GCGATGAGTG	GCAGGGCGGG	GCGTAAACGC	GTGGATCCGG	CTTACTAAAA	GCCAGATAAC
2521	AGTATGCGTA	TTTGCGCGCT	GATTTTTGCG	GTATAAGAAT	ATATACTGAT	ATGTATACCC-

FIGURE 24B

46/240

2581 GAAGTATGTC AAAAAGAGGT GTGCTATGAA GCAGCGTATT ACAGTGACAS TTGACAGCGA
2641 CAGCTATCAG TTGCTCAAGG CATATATGAT GTCAATATCT CCGGTCTGGT AAGCACAACC
2701 ATGCAGAATG AAGCCCGTCG TCTGCGTGCC GAACGCTGGA AAGCGGAAAA TCAGGAAGGG
2761 ATGGCTGAGG TCGCCCGGTT TATTGAAATG AACGGCTCTT TTGCTGACGA GAACAGGGAC
2821 TGGTGAAATG CAGTTTAAGG TTTACACCTA TAAAAGAGAG AGCCGTTATC GTCTGTTTGT
2881 GGATGTACAG AGTGATATTA TTGACACGCC CGGGCGACGG ATGGTGATCC CCCTGGCCAG
2941 TGCACGTCTG CTGTCAGATA AAGTCTCCCG TGAACCTTAC CCGGTGGTGC ATATCGGGGA
3001 TGAAAGCTGG CGCATGATGA CCACCGATAT GGCCAGTGTG CCGGTCTCCG TTATCGGGGA
3061 AGAAGTGGCT GATCTCAGCC ACCCGGAAAA TGACATCAA AACGCCATTA ACCTGATGTT
3121 CTGGGGAATA TAAATGTCAG GCTCCCTTAT ACACAGCCAG TCTGCAGGTC GACCATAGTG
3181 ACTGGATATG TTGTGTTTTA CAGTATTATG TAGTCTGTTT TTTATGCAAA ATCTAATTTA
3241 ATATATTGAT ATTTATATCA TTTTACGTTT CTCGTTTCAGC TTTCTTGATC AAAGTGGTGA
3301 TGGGGATCCT CTAGAGTCGA CCTGCAGTAA TCGTACAGGG TAGTACAAAT AAAAAGGCCA
3361 CGTCAGATGA CGTGCCCTTTT TTCTTGTCAG CAGTAAGCTT GGCTGTTTTG GCGGATGAGA
3421 TAAGATTTTC AGCCTGATAC AGATTAAATC AGAACGCAGA AGCGGTCTGA TAAACAGAA
3481 TTGCGCTGGC GGCAGTAGCC CGGTGGTCCC ACCTGACCCC ATGCCGAAC CAGAAGTGAA
3541 ACGCCGTAGC GCCGATGGTA GTGTGGGGTC TCCCCATGCG AGAGTAGGGA ACTGCCAGGC
3601 ATCAAATAAA ACGAAAGGCT CAGTCGAAAG ACTGGGCCTT TCGTTTTATC TGTGTTTGT
3661 CGGTGAACGC TCTCCTGAGT AGGACAAATC CGCCGGGAGC GGATTTGAAC GTTCCGAAGC
3721 AACGGCCCCG AGGGTGGCGG GCAGGACGCC CGCCATAAAC TGCCAGGCAT CAAATTAAGC
3781 AGAAGGCCAT CCTGACGGAT GGCCTTTTTG CGTTTCTACA AACTCTTTTT GTTTATTTTT
3841 CTAAATACAT TCAAATATGT ATCCGCTCAT GAGACAATAA CCCTGATAAA TGCTTCAATA
3901 ATATTGAAAA AGGAAGAGTA TGAGTATTCA ACATTTCCTG TCGCCCTTA TTCCCTTTTT
3961 TGCGGCATTT TGCCCTTCCTG TTTTGTCTCA CCCAGAAACG CTGGTGAAA3 TAAAAGATGC
4021 TGAAGATCAG TTGGGTGCAC GAGTGGGTTA CATCGAACTG GATCTCAACA GCGGTAAGAT
4081 CCTTGAGAGT TTTCCGCCCG AAGAACGTTT TCCAATGATG AGCATTTTA AAGTTCTGCT
4141 ATGTGGCGCG GTATTATCCC GTGTTGACGC CGGGCAAGAG CAACTCGGTC GCCGCATACA
4201 CTATTCTCAG AATGACTTGG TTGAGTACTC ACCAGTCACA GAAAAGCATC TTACGGATGG
4261 CATGACAGTA AGAGAATTAT GCAGTGCTGC CATAACCATG AGTGATAACA CTGCGGCCAA
4321 CTTACTTCTG ACAACGATCG GAGGACCGAA GGAGCTAACC GCTTTTTTGC ACAACATGGG
4381 GGATCATGTA ACTCGCCTTG ATCGTTGGGA ACCGGAGCTG AATGAAGCCA TACCAAACGA
4441 CGAGCGTGAC ACCACGATGC CTACAGCAAT GGCAACAACG TTGCGCAAA3 TATTAAGTGG
4501 CGAACTACTT ACTCTAGCTT CCCGGCAACA ATTAATAGAC TGGATGGAGG CGGATAAAGT
4561 TGCAGGACCA CTTCTGCGCT CGGCCCTTCC GGCTGGCTGG TTTATGCTG ATAAATCTGG
4621 AGCCGGTGAG CGTGGGTCTC GCGGTATCAT TGCAGCACTG GGGCCAGATG GTAAGCCCTC
4681 CCGTATCGTA GTTATCTACA CGACGGGGAG TCAGGCAACT ATGGATGAAC GAAATAGACA
4741 GATCGCTGAG ATAGGTGCCT CACTGATTAA GCATTGGTAA CTGTCAGACC AAGTTTACTC
4801 ATATATACTT TAGATTGATT TAAACTTCA TTTTAAATTT AAAAGGATCT AGGTGAAGAT
4861 CCTTTTTGAT AATCTCATGA CCAAAATCCC TTAACGTGAG TTTTCGTTCC ACTGAGCGTC
4921 AGACCCCGTA GAAAAGATCA AAGGATCTTC TTGAGATCCT TTTTTCTGTC GCGTAATCTG
4981 CTGCTTGCAA ACAAAAAAC CACCGCTACC AGCGGTGGTT TGTTTGCCGG ATCAAGAGCT
5041 ACCAACTCTT TTTCCGAAGG TAACTGGCTT CAGCAGAGCG CAGATACCAA ATACTGTCCT
5101 TCTAGTGTAG CCGTAGTTAG GCCACCACTT CAAGAACTCT GTAGCACCGC CTACATACCT
5161 CGCTCTGCTA ATCCTGTTAC CAGTGGCTGC TGCCAGTGGC GATAAGTCGT GTCTTACCGG
5221 GTTGGACTCA AGACGATAGT TACCGGATAA GGCGCAGCGG TCGGGCTGAA CGGGGGGTTT
5281 GTGCACACAG CCCAGCTTGG AGCGAACGAC CTACACCGAA CTGAGATACC TACAGCGTGA
5341 GCTATGAGAA AGCGCCACGC TTCCCGAAGG GAGAAAGGCG GACAGGTATC CGGTAAGCGG
5401 CAGGGTCGGA ACAGGAGAGC GCACGAGGGA GCTTCCAGGG GGAACGCTT GGTATCTTTA
5461 TAGTCCTGTC GGGTTTCGCC ACCTCTGACT TGAGCGTCGA TTTTTGTGAT GCTCGTCAGG
5521 GGGGCGGAGC CTATGGAAAA ACGCCAGCAA CGCGGCCTTT TTACGGTTCC TGGCCTTTTG
5581 CTGGCCTTTT GCTCACATGT TCTTTCCTGC GTTATCCCCT GATTCTGTGG ATAACCGTAT
5641 TACCGCCTTT GAGTGAGCTG ATACCGCTCG CCGCAGCCGA ACGACCGAGC GCAGCGAGTC
5701 AGTGAGCGAG GAAGCGGAAG AGCGCTGAT GCGGTATTTT CTCCTTACGC ATCTGTGCGG
5761 TATTTACAC CACATAATTT TGTTAAATTT CGCGTTAAAT TTTTGTTAAA TCAGCTCATT
5821 TTTTAAACCA TAGGCCGAAA TCGGCAAAAT CCCTTATAAA TCAAAAGAAT AGACCGAGAT
5881 AGGGTTGAGT GTTGTTCAG TTTGGAACAA GAGTCCACTA TTAAAGAACG TGGAATCCAA
5941 CGTCAAAGGG CGAAAAACCG TCTATCAGGG CGATGGCCCA CTACGTGAAC CATCACCTTA
6001 ATCAAGTTTT TTGGGGTCGA GGTGCCGTAA AGCACTAAAT CGGAACCTTA AAGGGAGCCC-

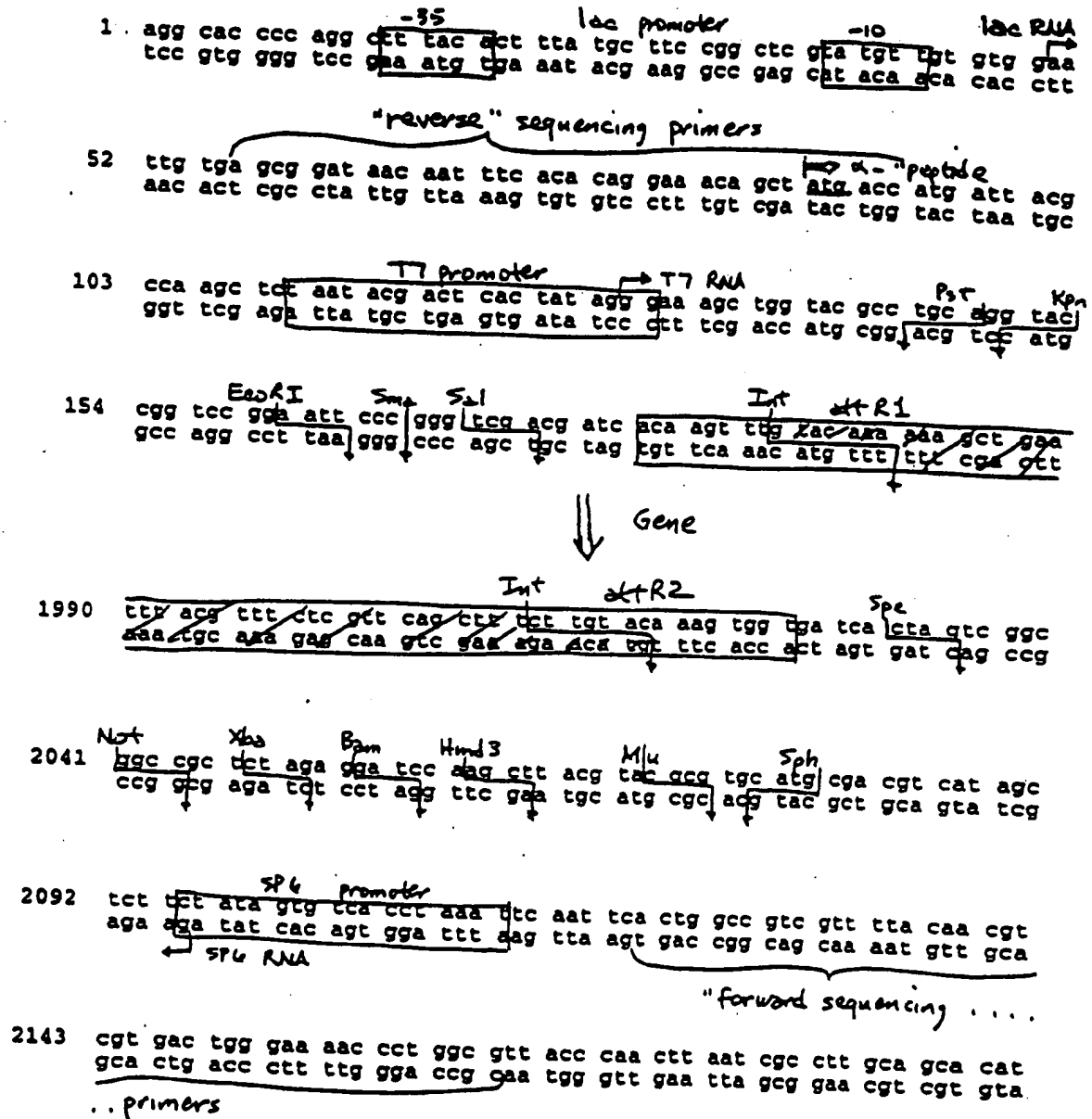
FIGURE 24C

6061 CCGATTTAGA GCTTGACGGG GAAAGCCGGC GAACGTGGCG AGAAAGGAAG GGAAGAAAGC
6121 GAAAGGAGCG GGCGCTAGGG CGCTGGCAAG TGTAGCGGTC ACGCTGCGCG TAACCACCAC
6181 ACCCGCCGCG CTTAATGCGC CGCTACAGGG CGCGTCCATT CGCCATTTCAG GCTGCTATGG
6241 TGCACTCTCA GTACAATCTG CTCTGATGCC GCATAGTTAA GCCAGTATAC ACTCCGCTAT
6301 CGCTACGTGA CTGGGTCAATG GCTGCGCCCC GACACCCGCC AACACCCGCT GACGCGCCCT
6361 GACGGGCTTG TCTGCTCCCG GCATCCGCTT ACAGACAAGC TGTGACCGTC TCCGGGAGCT
6421 GCATGTGTCA GAGGTTTTCA CCGTCATCAC CGAAACGCGC GAGGCAGCAG ATCAATTTCG
6481 GCGCGAAGGC GAAGCGGCAT GCATTACGT TGACACCATC GAATGGTGCA AAACCTTTCG
6541 CGGTATGGCA TGATAGCGCC CGGAAGAGAG TCAATTCAGG GTGGTGAATG TGAAACCAGT
6601 AACGTTATAC GATGTCGCAG AGTATGCCGG TGTCTCTTAT CAGACCGTTT CCCGCGTGGT
6661 GAACCAGGCC AGCCACGTTT CTGCGAAAAC GCGGGAAAAA GTGGAAGCGG CGATGGCGGA
6721 GCTGAATTAC ATTCCCAACC GCGTGGCACA ACAACTGGCG GGCAAACAGT CGTTGCTGAT
6781 TGGCGTTGCC ACCTCCAGTC TGGCCCTGCA CGCGCCGTCG CAAATTGTCG CGGCGATTAA
6841 ATCTCGCGCC GATCAACTGG GTGCCAGCGT GGTGGTGTG ATGGTAGAAC GAAGCGGCGT
6901 CGAAGCCTGT AAAGCGGCGG TGCACAATCT TCTCGCGCAA CGCGTCAGTN GGGCTGATCA
6961 TTAA

48/240

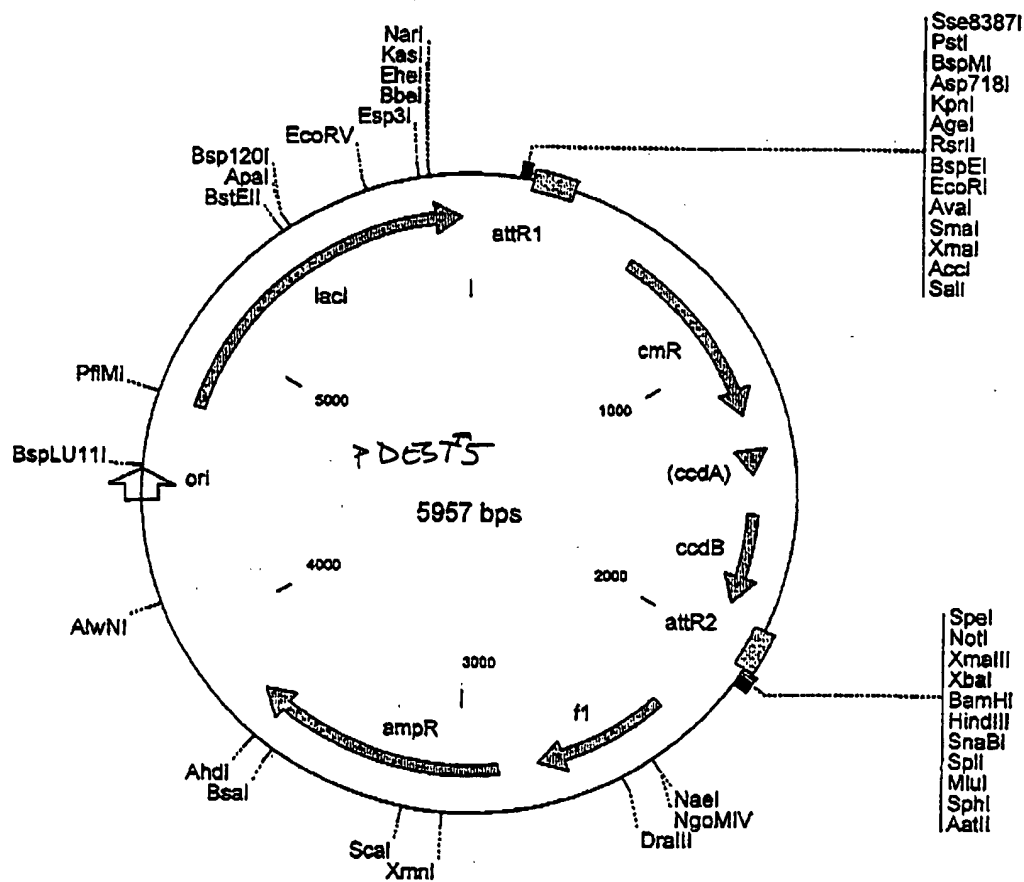
Figure 25A

pDEST5

pSPORT '+' (for sequencing, probes,
phagemid)

49/260

Figure 25B γ DEST5 (cont'd)



50/240

pDEST5 5957 bp

<u>Location (Base Nos.)</u>		<u>Gene Encoded</u>
305..181		attR1
555..1214		CmR
1334..1418		inactivated ccdA
1556..1861		ccdB
1902..2026		attR2
2278..2733		f1 (f1 intergenic region)
2865..3722		ampR
5378..5538		ori
4756..5922		lacI

1	AGGCACCCCA	GGCTTTACAC	TTTATGCTTC	CGGCTCGTAT	GTTGTGTGGA	ATTGTGAGCG
61	GATAACAATT	TCACACAGGA	AACAGCTATG	ACCATGATTA	CGCCAAGCTC	TAATACGACT
121	CACTATAGGG	AAAGCTGGTA	CGCCTGCAGG	TACCGGTCCG	GAATTCCCGG	GTCGACGATC
181	ACAAGTTTGT	ACAAAAAAGC	TGAACGAGAA	ACGTAAATG	ATATAAATAT	CAATATATTA
241	AATTAGATTT	TGCATAAAAA	ACAGACTACA	TAATACTGTA	AAACACAACA	TATCCAGTCA
301	CTATGGCGGC	CGCTAAGTTG	GCAGCATCAC	CCGACGCACT	TTGCGCCGAA	TAAATACCTG
361	TGACGGAAGA	TCACTTCGCA	GAATAAATAA	ATCCTGGTGT	CCCTGTTGAT	ACCGGGAAGC
421	CCTGGGCCAA	CTTTTGGCGA	AAATGAGACG	TTGATCGGCA	CGTAAGAGGT	TCCAACCTTC
481	ACCATAATGA	AATAAGATCA	CTACCGGGCG	TATTTTTTGA	GTTATCGAGA	TTTTTCAGGAG
541	CTAAGGAAGC	TAAAATGGAG	AAAAAATCA	CTGGATATAC	CACCGTTGAT	ATATCCCAAT
601	GGCATCGTAA	AGAACATTTT	GAGGCATTTT	AGTCAGTTGC	TCAATGTACC	TATAACCAGA
661	CGGTTTCAGCT	GGATATTACG	GCCTTTTTTA	AGACCGTAAA	GAAAAATAAG	CACAAGTTTT
721	ATCCGGCCTT	TATTCACATT	CTTGCCCGCC	TGATGAATGC	TCATCCGGAA	TTCCGTATGG
781	CAATGAAAGA	CGGTGAGCTG	GTGATATGGG	ATAGTGTTCA	CCCTTGTTAC	ACCGTTTTTC
841	ATGAGCAAAC	TGAAACGTTT	TCATCGCTCT	GGAGTGAATA	CCACGACGAT	TTCCGGCAGT
901	TTCTACACAT	ATATTGCAAA	GATGTGGCGT	GTTACGGTGA	AAACCTGGCC	TATTTCCCTA
961	AAGGGTTTTAT	TGAGAATATG	TTTTTCGTCT	CAGCCAATCC	CTGGGTGAGT	TTCACCAGTT
1021	TTGATTTAAA	CGTGGCCAAT	ATGGACAAC	TCTTCGCCCC	CGTTTTTACC	ATGGGCAAT
1081	ATTATACGCA	AGGCGACAAG	GTGCTGATGC	CGCTGGCGAT	TCAGGTTTCT	CATGCCGTCT
1141	GTGATGGCTT	CCATGTCGGC	AGAATGCTTA	ATGAATTACA	ACAGTACTGC	GATGAGTGGC
1201	AGGGCGGGGC	GTAAACGCGT	GGATCCGGCT	TACTAAAAGC	CAGATAACAG	TATGCGTATT
1261	TGCGCGCTGA	TTTTTGCGGT	ATAAGAATAT	ATACTGATAT	GTATACCCGA	AGTATGTCAA
1321	AAAGAGGTGT	GCTATGAAGC	AGCGTATTAC	AGTGACAGTT	GACAGCGACA	GCTATCAGTT
1381	GCTCAAGGCA	TATATGATGT	CAATATCTCC	GGTCTGGTAA	GCACAACCAT	GCAGAATGAA
1441	GCCCGTCTGC	TGCGTGCCGA	ACGCTGGAAG	GCGGAAATC	AGGAAGGGAT	GGCTGAGGTC
1501	GCCCGGTTTA	TTGAAATGAA	CGGCTCTTTT	GCTGACGAGA	ACAGGGACTG	GTGAAATGCA
1561	GTTTAAGGTT	TACACCTATA	AAAGAGAGAG	CCGTTATCGT	CTGTTTGTGG	ATGTACAGAG
1621	TGATATTATT	GACACGCCCC	GGCGACGGAT	GGTGATCCCC	CTGGCCAGTG	CACGTCGTCT
1681	GTCAGATAAA	GTCTCCCGTG	AACTTTACCC	GGTGGTGAT	ATCGGGGATG	AAAGCTGGCG
1741	CATGATGACC	ACCGATATGG	CCAGTGTGCC	GGTCTCCGTT	ATCGGGGAAG	AAGTGGCTGA
1801	TCTCAGCCAC	CGCGAAATG	ACATCAAAAA	CGCCATTAAC	CTGATGTTCT	GGGGAATATA
1861	AATGTCAGGC	TCCCTTATAC	ACAGCCAGTC	TGCAGGTCGA	CCATAGTGAC	TGGATATGTT
1921	GTGTTTTACA	GTATTATGTA	GTCTGTTTTT	TATGCAAAAT	CTAATTTAAT	ATATTGATAT
1981	TTATATCATT	TTACGTTTCT	CGTTCAGCTT	TCTTGTAACA	AGTGGTGATC	ACTAGTCGGC
2041	GGCCGCTCTA	GAGGATCCAA	GCTTACGTAC	GCGTGATGAC	GACGTCATAG	CTCTTCTATA
2101	GTGTACACCTA	AATTCAATTC	ACTGGCCGTC	GTTTTTACAAC	GTCGTGACTG	GGAAAACCTT
2161	GGCGTTACCC	AACTTAATCG	CCTTGACGCA	CATCCCCCTT	TCGCCAGCTG	GCGTAATAGC
2221	GAAGAGGCCC	GCACCGATCG	CCCTTCCCAA	CAGTTGCGCA	GCCTGAATGG	CGAATGGACG
2281	CGCCCTGTAG	CGGCGCATTA	AGCGCGGCGG	GTGTGGTGGT	TACGCGCAGC	GTGACCGCTA
2341	CACCTTGCCAG	CGCCCTAGCG	CCCGCTCCTT	TCGCTTTCTT	CCCTTCTTTT	CTCGCCACGT
2401	TGCGCGGCTT	TCCCCGTCAA	GCTCTAAATC	GGGGGCTCCC	TTTAGGGTTC	CGATTTAGTG
2461	CTTTACGGCA	CCTCGACCCC	AAAAAACTTG	ATTAGGGTGA	TGGTTCACGT	AGTGGGCCAT
2521	CGCCCTGATA	GACGGTTTTT	CGCCCTTTGA	CGTTGGAGTC	CACGTTCTTT	AATAGTGGAC
2581	TCTTGTTCCA	AACTGGAACA	ACACTCAACC	CTATCTCGGT	CTATTCTTTT	GATTTATAAG-

FIGURE 25C

51/240

2641 GGATTTTGCC GATTTCGGCC TATTGGTTAA AAAATGAGCT GATTTAACAA AAATTTAACG
2701 CGAATTTTAA CAAAATATTA ACGTTTACAA TTTCAGGTGG CACTTTTCGG GGAAATGTGC
2761 GCGGAACCCC TATTTGTTTA TTTTCTAAA TACATTCAAA TATGTATCCG CTCATGAGAC
2821 AATAACCCCTG ATAAATGCTT CAATAATATT GAAAAAGGAA GAGTATGAGT ATTCAACATT
2881 TCCGTGTGCG CTTTATTCCT TTTTGTGCGG CATTTTGCCT TCCTGTTTTT GCTCACCAG
2941 AAACGCTGGT GAAAGTAAAA GATGCTGAAG ATCAGTTGGG TGCACGAGTG GGTACATCG
3001 AACTGGATCT CAACAGCGGT AAGATCCTTG AGAGTTTTCG CCCCAGAGAA CGTTTTCCAA
3061 TGATGAGCAC TTTTAAAGTT CTGCTATGTG GCGCGGTATT ATCCCGTATT GACGCGGGC
3121 AAGAGCAACT CGGTCGCCG ATACACTATT CTCAGAATGA CTTGGTTGAG TACTCACCAG
3181 TCACAGAAAA GCATCTTACG GATGGCATGA CAGTAAGAGA ATTATGCAGT GCTGCCATAA
3241 CCATGAGTGA TAACACTGCG GCCAACTTAC TTCTGACAAC GATCGGAGGA CCGAAGGAGC
3301 TAACCGCTTT TTTGCACAAC ATGGGGGATC ATGTAACTCG CTTTGATCGT TGGGAACCGG
3361 AGCTGAATGA AGCCATACCA AACGACGAGC GTGACACCAC GATGCCTGTA GCAATGGCAA
3421 CAACGTTGCG CAAACTATTA ACTGGCGAAC TACTTACTCT AGCTTCCCGG CAACAATTAA
3481 TAGACTGGAT GGAGGCGGAT AAAGTTGCAG GACCACTTCT GCGCTCGGCC CTTCCGGCTG
3541 GCTGGTTTAT TGCTGATAAA TCTGGAGCCG GTGAGCGTGG GTCTCGCGGT ATCATTTGCAG
3601 CACTGGGGCC AGATGGTAAG CCTCCCCTA TCGTAGTTAT CTACACGAGC GGGAGTCAGG
3661 CAACTATGGA TGAACGAAAT AGACAGATCG CTGAGATAGG TGCCTCACTG ATTAAGCATT
3721 GGTAAGTGT AGACCAAGTT TACTCATATA TACTTTAGAT TGATTTAAAA CTTTCATTTTT
3781 AATTTAAAG GATCTAGGTG AAGATCCTTT TTGATAATCT CATGACCAA ATCCCTTAAC
3841 GTGAGTTTTT GTTCCACTGA GCGTCAGACC CCGTAGAAAA GATCAAAGGA TCTTCTTGAG
3901 ATCCTTTTTT TCTGCGCGTA ATCTGCTGCT TGCAAAACAA AAAACCACCG CTACCAGCGG
3961 TGGTTTGTGTT GCCGGATCAA GAGCTACCAA CTCTTTTTCC GAAGGTAAGT GGCTTCAGCA
4021 GAGCGCAGAT ACCAAATACT GTCCTTCTAG TGTAGCCGTA GTTAGGCCAC CACTTCAAGA
4081 ACTCTGTAGC ACCGCCTACA TACCTCGCTC TGCTAATCCT GTTACCAAGT GCTGCTGCCA
4141 GTGGCGATAA GTCGTGTCTT ACCGGGTTGG ACTCAAGACG ATAGTTACCG GATAAGGCGC
4201 AGCGGTGCGG CTGAACGGGG GGTTCGTGCA CACAGCCAG CTTGGAGCGA ACGACCTACA
4261 CCGAACTGAG ATACCTACAG CGTGAGCATT GAGAAAGCGC CACGCTTCCC GAAGGAGAA
4321 AGGCGGACAG GTATCCGGTA AGCGGCAGGG TCGGAACAGG AGAGCGCAGC AGGGAGCTTC
4381 CAGGGGGAAA CGCCTGGTAT CTTTATAGTC CTGTGCGGTT TCGCCACCTC TGACTTGAGC
4441 GTCGATTTTT GTGATGCTCG TCAGGGGGGC GGAGCCTATG GAAAAACGCC AGCAACGCGG
4501 CCTTTTTTACG GTTCTTGCC TTTTGCTGGC CTTTTGCTCA CATGTTCTTT CCTGCGTTAT
4561 CCCCTGATTC TGTGGATAAC CGTATTACCG CTTTGAGTG AGCTGATACC GCTCGCCGCA
4621 GCCGAACGAC CGAGCGCAGC GAGTCAGTGA GCGAGGAAGC GGAAGAGCGC CCAATACGCA
4681 AACCGCCTCT CCCCGCGCTT TGGCCGATT ATTAATGCAG AGCTTGCAAT TCGCGCGCGA
4741 AGGCGAAGCG GCATTTACGT TGACACCATC GAATGGCGCA AAACCTTTTC CGGTATGGCA
4801 TGATAGCGCC CGGAAGAGAG TCAATTCAGG GTGGTGAATG TGAAACCACT AACGTTATAC
4861 GATGTCGAG AGTATGCCGG TGTCTCTTAT CAGACCGTTT CCCGCGTGGT GAACAGGCC
4921 AGCCACGTTT CTGCGAAAAC GCGGAAAAA GTGGAAGCGG CGATGGCGGA GCTGAATTAC
4981 ATTCCCAACC GCGTGGCACA ACAACTGGCG GGCAAACAGT CGTTGCTGAT TGGCGTTGCC
5041 ACCTCCAGTC TGGCCCTGCA CGCGCCGTCG CAAATTGTCT GCGCGATTAA ATCTCGCGCC
5101 GATCAACTGG GTGCCAGCGT GGTGGTGTCT ATGGTAGAAC GAAGCGGCGT CGAAGCCTGT
5161 AAAGCGGCGG TGACAACTCT TCTCGCGCAA CGGGTCAGTG GGCTGATCAT TAACATCCG
5221 CTGGATGACC AGGATGCCAT TGCTGTGGAA GCTGCCTGCA CTAATGTTCC GCGGTTATTT
5281 CTTGATGTCT CTGACCAGAC ACCCATCAAC AGTATTATTT TCTCCCATGA AGACGGTACG
5341 CGACTGGGCG TGGAGCATCT GGTGCGATTG GGTACCCAGC AAATCGCGCT GTTAGCGGGC
5401 CCATTAAGTT CTGTCTCGGC GCGTCTGCGT CTGGCTGGCT GGCATAAATA TCTCACTCGC
5461 AATCAAATTC AGCCGATAGC GGAACGGGAA GCGCACTGGA GTGCCATGTC CGGTTTTCAA
5521 CAAACCATGC AAATGCTGAA TGAGGCAATC GTTCCCACTG CGATGCTGGT TGCCAACGAT
5581 CAGATGGCGC TGGGCGCAAT GCGCGCCATT ACCGAGTCCG GGCTGCGCGT TGGTGGGAT
5641 ATCTCGGTAG TGGGATACGA CGATACCGAA GACAGCTCAT GTTATATCCC GCCGTCAACC
5701 ACCATCAAAC AGGATTTTCG CCTGCTGGGG CAAACCAGCG TGGACCGCTT GCTGCAACTC
5761 TCTCAGGGCC AGGCGGTGAA GGGCAATCAG CTGTTGCCCG TCTCACTGGT GAAAAGAAAA
5821 ACCACCTGG CGCCCAATAC GCAAACCGCC TCTCCCCGCG CGTTGGCCGA TTCAATTAATG
5881 CAGCTGGCAC GACAGGTTTC CCGACTGGAA AGCGGCGAGT GAGCGCAACG CAATTAATGT
5941 GAGTTAGCTC ACTCAT

FIGURE 25D

Figure 26A

pDEST6

pSPORT "..."
(opposite strand)

"forward" sequencing primers

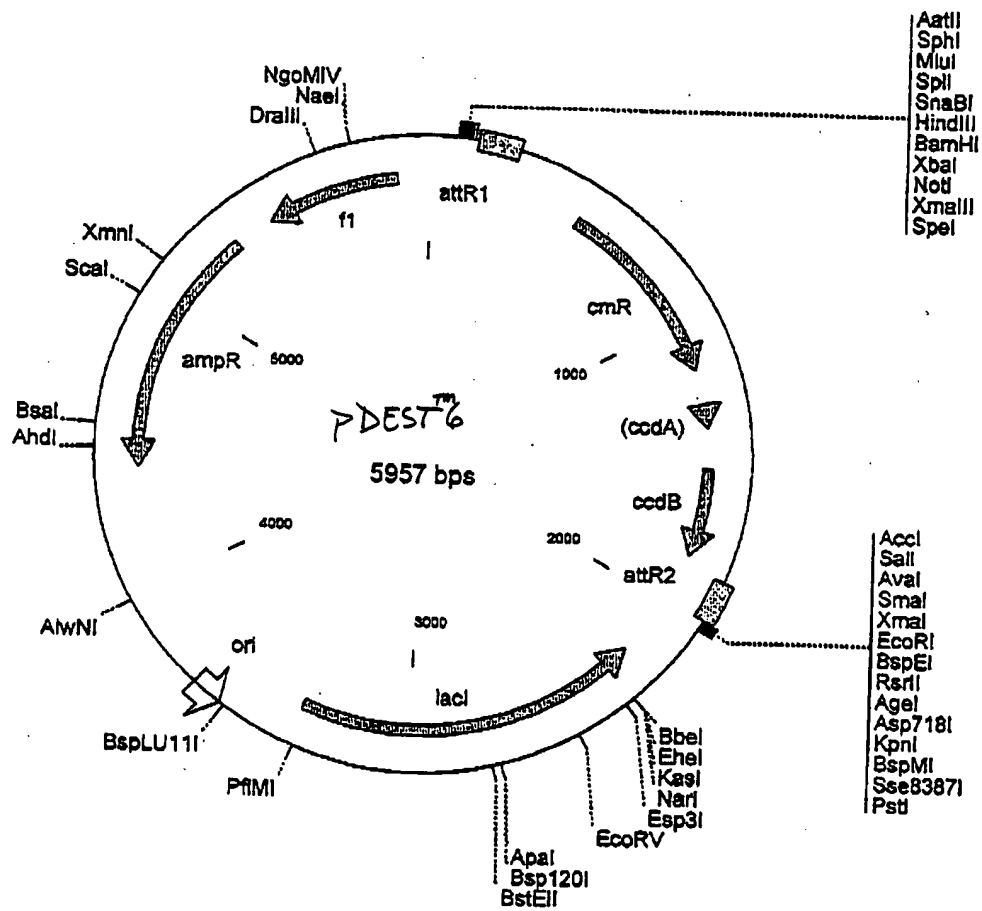
- 1 taa tgc cag ggt ttt ccc agt cac gac gtt gta aaa cga cgg cca gtc aat
att gcg gtc cca aaa ggg tca gtg ctg caa cat ttt gct gcc ggt cac tta
- 52 ^{SP6 promoter} tga att tag gtg aca cta tag aag agc tat gac gtc gca tgt acg cgt acg
act taa atc cac tgt gat atc ttc tgc ata ctg cag cgt acg tgc gca tgc
- 103 ^{Hind3} tga gct tgg atc ctc tag agc ^{Bam} agc cgc cga cta gtg atc ^{Xba} ata gta atc ^{Not} ata gta atc ^{Spe} ata gta atc ^{Int} ata gta atc
att cga acc tag gag atc tgc ccg ggc gct gat gac tag tgt tca aac atg
- 154 ~~aaa daa gct gaa cga gaa acg taa aat gar ata aat atc aat ata tta aat
ttt tat cga ctt gct ctt tga att tta cta tat tca tag tta tat aat tca~~
- ↓ Gene
- 1939 ~~tat tta tat tat ttt acg att ctc gtt tag ctt gct tgt aca aag tgg tga
ata aat ata gta aaa tgc aaa gag aaa gtc gaa aga aca tgc ttc acc att~~
- 1990 ^{Sol} tgc tgc acc cgg ^{Spy} gaa ttc cgg acc ^{EcoRI} ggt acg tgc ggg cgt acc agc ttt ccc
agc agc tgg gcc ctt aag gcc tgg gca tgg acg tcc gca tgg tgc aaa ggg
- T7 RNA
- 2041 tat agt gag tgc tat tag agc ttg gcg taa tca tgg tca tag ctg ttt cct
ata tca ctc agc ata atc tgc aac cgc att agt acc agt atc gac aaa gga
- T7 promoter α-peptide "reverse .."
- 2092 gtg tga aat tgt tat cgg ctc aca att cca cac ⁻¹⁰ aac ata ^{lac promoter} cga gct gga agc
cac act tta aca ata ggc gag tgt taa ggt gtg ttg tat gct cgg cct tgc
- ... sequencing primers lac RNA
- 2143 ⁻³⁵ ata aag tgt aaa gcc tgg ggt gcc taa tga gtg agc taa ctc aca tta att
tat ttc aca ttt cgg acc cca cgg att act cac tgc att gag tgt aat taa

53/240

Figure 26B

PDEST6

(cont'd)



54/240

pDEST6 5957 bp

<u>Location (Base Nos.)</u>		<u>Gene Encoded</u>
266..142		attR1
516..1175		CmR
1295..1379		inactivated ccdA
1517..1822		ccdB
1863..1987		attR2
2203..3369		lacI
4403..5260		ampR
5392..5847		f1 (f1 intergenic region)
1	TAACGCCAGG GTTTTCCCAG TCACGACGTT GTAAAAACGAC GGCCAGTGAA TTGAATTTAG	
61	GTGACACTAT AGAAGAGCTA TGACGTCGCA TGCACGCGTA CGTAAGCTTG GATCCTCTAG	
121	AGCGGCCGCC GACTAGTGAT CACAAGTTTG TACAAAAAAG CTGAACGAGA AACGTAAAAT	
181	GATATAAATA TCAATATATT AAATTAGATT TTGCATAAAA AACAGACTAC ATAATACTGT	
241	AAACACAAAC ATATCCAGTC ACTATGGCGG CCGCTAAGTT GGCAGCATCA CCCGACGCAC	
301	TTTGGCGCGA ATAAATACCT GTGACGGAAG ATCACTTCGC AGAATAAATA AATCCTGGTG	
361	TCCCTGTTGA TACCGGGGAG CCCTGGGCCA ACTTTTGGCG AAAATGAGAC GTTGATCGGC	
421	ACGTAAGAGG TTCCAACTTT CACCATAATG AAATAAGATC ACTACCGGGC GTATTTTTTG	
481	AGTTATCGAG ATTTTCAGGA GCTAAGGAAG CTAAAATGGA GAAAAAATC ACTGGATATA	
541	CCACCGTTGA TATATCCCAA TGGCATCGTA AAGAACATTT TGAGGCATTT CAGTCAGTTG	
601	CTCAATGTAC CTATAACCAG ACCGTTTCAGC TGGATATTAC GGCTTTTTTA AAGACCGTAA	
661	AGAAAAATAA GCACAAGTTT TATCCGGCCT TTATTCACAT TCTTGCCCGC CTGATGAATG	
721	CTCATCCGGA ATTCCGTATG GCAATGAAAG ACGGTGAGCT GGTGATATGG GATAGTGTTT	
781	ACCCTTGTGA CACCGTTTTT CATGAGCAAA CTGAAACGTT TTCATCGCTC TGGAGTGAAT	
841	ACCACGACGA TTTCCGGCAG TTTCTACACA TATATTCGCA AGATGTGGCG TGTACGGTG	
901	AAAACCTGGC CTATTTCCCT AAAGGGTTTA TTGAGAATAT GTTTTTCGTC TCAGCCAATC	
961	CCTGGGTGAG TTTCAACAGT TTTGATTAA ACGTGGCCAA TATGGACAAC TTCTTCGCCC	
1021	CCGTTTTTAC CATGGGCAAA TATTATACGC AAGGCGACAA GGTGCTGATG CCGCTGGCGA	
1081	TTCAGGTTCA TCATGCCGTC TGTGATGGCT TCCATGTCCG CAGAAATGCTT AATGAATTAC	
1141	AACAGTACTG CGATGAGTGG CAGGGCGGGG CGTAAACGCG TGATCCGGC TTAATAAAG	
1201	CCAGATAACA GTATGCGTAT TTGCGCGCTG ATTTTTCGGG TATAAGAATA TATACTGATA	
1261	TGTATACCCG AAGTATGTCA AAAAGAGGTG TGCTATGAAG CAGCGTATTA CAGTGACAGT	
1321	TGACAGCGAC AGCTATCAGT TGCTCAAGGC ATATATGATG TCAATATCTC CGGTCTGGTA	
1381	AGCACAACCA TGCAGAAATG AGCCCGTCGT CTGCGTGCCG AACGCTGGAA AGCGGAAAAT	
1441	CAGGAAGGGA TGGCTGAGGT CGCCCGGTTT ATTGAAATGA ACGGCTCTTT TGCTGACGAG	
1501	AACAGGGACT GGTGAAATGC AGTTTAAAGT TTACACCTAT AAAAGAGAGA GCCGTTATCG	
1561	TCTGTTTGTG GATGTACAGA GTGATATTAT TGACACGCCC GGGCGACGGA TGGTGATCCC	
1621	CCTGGCCAGT GCACGTCTGC TGTCAGATAA AGTCTCCCGT GAACTTTACC CGGTGGTGCA	
1681	TATCGGGGAT GAAAGCTGGC GCATGATGAC CACCGATATG GCCAGTGTGC CGGTCTCCGT	
1741	TATCGGGGAA GAAGTGGCTG ATCTCAGCCA CCGCGAAAAT GACATCAAAA ACGCCATTAA	
1801	CCTGATGTTC TGGGGAATAT AAATGTCAGG CTCCCTTATA CACAGCCAGT CTGCAGGTCG	
1861	ACCATAGTGA CTGGATATGT TGTGTTTTAC AGTATTATGT AGTCTGTTTT TTATGCAAAA	
1921	TCTAATTTAA TATATTGATA TTTATATCAT TTTACGTTT TCCTTCAGCT TTCTTGATCA	
1981	AAGTGGTGAT CGTCGACCCG GGAATTCGG ACCGGTACCT GCAGGCGTAC CAGCTTTCCC	
2041	TATAGTGAAT CGTATTAGAG CTTGGCGTAA TCATGGTCAT AGCTGTTTCC TGTGTGAAAT	
2101	TGTTATCCGC TCACAATTCC ACACAACATA CGAGCCGGAA GCATAAAGTG TAAAGCCTGG	
2161	GGTGCCCTAAT GAGTGAGCTA ACTCACATTA ATTGCGTTGC GCTCACTGCC CGCTTTCCAG	
2221	TCGGGAAACC TGTCGTGCCA GCTGCATTAA TGAATCGGCC AACGCGCGGG GAGAGGCGGT	
2281	TTGCGTATTG GGCGCCAGGG TGGTTTTTCT TTTACCAAGT GAGACGGGCA ACAGCTGATT	
2341	GCCCTTCACC GCCTGGCCCT GAGAGAGTTG CAGCAAGCGG TCCACGCTGG TTTGCCCCAG	
2401	CAGGCGAAAA TCCTGTTTGA TGGTGTTTGA CGGCGGGATA TAACATGAGC TGTCTTCGGT	
2461	ATCGTCGTAT CCCACTACCG AGATATCCGC ACCAACCGGC AGCCCGGACT CGGTAATGGC	
2521	GCGCATTGCG CCCAGCGCCA TCTGATCGTT GGCAACCAGC ATCGCAGTGG GAACGATGCC	
2581	CTCATTCAGC ATTTGTCATG TTTGTTGAAA ACCGGACATG GCACTCCAGT CGCCTTCCCG	
2641	TTCCGCTATC GGCTGAATTT GATTGCGAGT GAGATATTTA TGCCAGCCAG CCAGACGCAG-	

FIGURE 26C

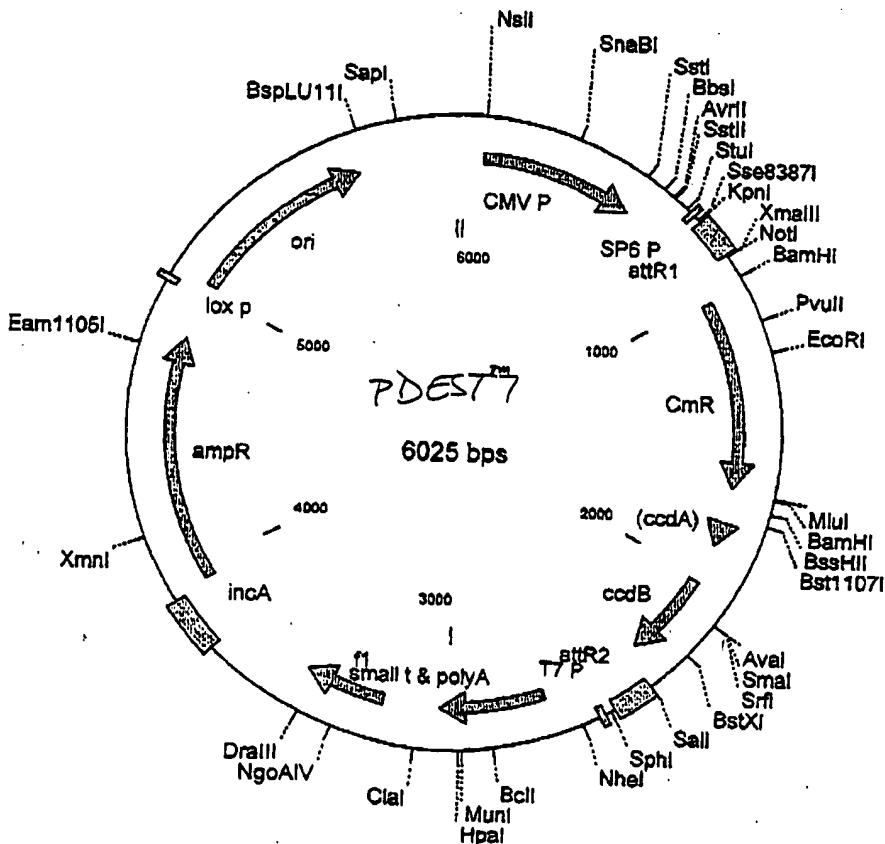
2701 ACGCGCCGAG ACAGAACTTA ATGGGCCCGC TAACAGCGCG ATTTGCTGGT GACCCAATGC
2761 GACCAGATGC TCCACGCCCA GTCGCGTACC GTCTTCATGG GAGAAAATAA TACTGTTGAT
2821 GGGTGTCTGG TCAGAGACAT CAAGAAATAA CGCCGGAACA TTAGTGACAG CAGCTTCCAC
2881 AGCAATGGCA TCCTGGTCAT CCAGCGGATA GTTAATGATC AGCCCACTGA CCCGTTGCGC
2941 GAGAAGATTG TGCACCGCCG CTTTACAGGC TTGACGCGG CTTCGTTCTA CCATCGACAC
3001 CACCACGCTG GCACCCAGTT GATCGGCGCG AGATTTAATC GCCGCGACAA TTTGCGACGG
3061 CGCGTGACGG GCCAGACTGG AGGTGGCAAC GCCAATCAGC AACGACTGTT TGCCCGCCAG
3121 TTGTTGTGCC ACGCGGTTGG GAATGTAATT CAGCTCCGCC ATCGCCGCTT CCACTTTTTC
3181 CCGCGTTTTT GCAGAAACGT GGTGGCCTG GTTACCACG CGGAAACGG TCTGATAAGA
3241 GACACCGGCA TACTCTGCGA CATCGTATAA CGTTACTGGT TTCACATTCA CCACCTGAA
3301 TTGACTCTCT TCCGGGCGCT ATCATGCCAT ACCGCGAAAG GTTTTGCGCC ATTCGATGGT
3361 GTCAACGTAA ATGCCGCTTC GCCTTCGCGC GCGAATTGCA AGCTCTGCAT TAATGAATCG
3421 GCCAACGCGC GGGGAGAGGC GGTTCGCTA TTGGGCGCTC TTCCGCTTCC TCGCTCACTG
3481 ACTCGCTGCG CTCGGTCTGT CCGCTGCGGC GAGCGGTATC AGCTCACTCA AAGGCGGTAA
3541 TACGGTTATC CACAGAATCA GGGGATAACG CAGGAAAGAA CATGTGAGCA AAAGGCCAGC
3601 AAAAGGCCAG GAACCGTAAA AAGGCCGCGT TGCTGGCGTT TTTCCATAGG CTCCGCCCCC
3661 CTGACGAGCA TCACAAAAAT CGACGCTCAA GTCAGAGGTG GCGAAACCCG ACAGGACTAT
3721 AAAGATACCA GCGTTTTCCC CCTGGAAGCT CCTCTGTGCG CTCTCTGTT CCGACCTGCG
3781 CGCTTACCGG ATACCTGTCC GCCTTTCTCC CTTGCGGAAG CGTGGCGCTT TCTCAATGCT
3841 CACGCTGTAG GTATCTCAGT TCGGTGTAGG TCGTTCGCTC CAAGCTGGGC TGTGTGACAG
3901 AACCCCCCGT TCAGCCCGAC CGCTGCGCCT TATCCGGTAA CTATCGTCTT GAGTCCAACC
3961 CGGTAAGACA CGACTTATCG CCACTGGCAG CAGCCACTGG TAACAGGATT AGCAGAGCGA
4021 GGTATGTAGG CGGTGCTACA GAGTCTTGA AGTGGTGGCC TAACTACGGC TACACTAGAA
4081 GGACAGTATT TGGTATCTGC GCTCTGCTGA AGCCAGTTAC CTTGCGAAAA AGAGTTGGTA
4141 GCTCTTGATC CGGCAACAA ACCACCGCTG GTAGCGGTGG TTTTTTTGTT TGCAAGCAGC
4201 AGATTACGCG CAGAAAAAAA GGATCTCAAG AAGATCCTTT GATCTTTTCT ACGGGGTCTG
4261 ACGCTCAGTG GAACGAAAC TCACGTTAAG GGATTTTGGT CATGAGATTA TCAAAAAGGA
4321 TCTTCACTTA GATCCTTTTA AATTAAAAAT GAAGTTTTAA ATCAATCTAA AGTATATATG
4381 AGTAACTTGG GTCTGACAGT TACCAATGCT TAATCAGTGA GGCACCTATC TCAGCGATCT
4441 GTCTATTTTC TTCTATCCATA GTTGCTGAC TCCCGCTCGT GTAGATAACT ACGATACGGG
4501 AGGGCTTACC ATCTGGCCCC AGTGTGCAA TGATACCGCG AGACCCACGC TCACCGGCTC
4561 CAGATTATAT AGCAATAAAC CAGCCAGCCG GAAGGGCCGA GCGCAGAAAT GGTCTGCAA
4621 CTTTATCCGC CTCCATCCAG TCTATTAAAT GTTGCCGGGA AGCTAGAGTA AGTAGTTCCG
4681 CAGTTAATAG TTTGCGCAAC GTTGTGGCA TTGCTACAGG CATCGTGGTG TCACGCTCGT
4741 CGTTTGGTAT GGCTTCATTC AGCTCCGGT CCCAACGATC AAGGCGAGTT ACATGATCCC
4801 CCATGTTGTG CAAAAAAGCG GTTAGCTCCT TCGGTCTCTC GATCGTTGTC AGAAGTAAGT
4861 TGGCCGCACT GTTATCACTC ATGGTTATGG CAGCACTGCA TAATTCTCTT ACTGTCTATG
4921 CATCCGTAAG ATGCTTTTCT GTGACTGGTG AGTACTCAAC CAAGTCATTG TGAGAATAGT
4981 GTATGCGGCG ACCGAGTTGC TCTTGCCCGG CGTCAATACG GGATAATACC GCGCCACATA
5041 GCAGAACTTT AAAAGTGCTC ATCATTGGAA AACGTTCTTC GGGGCGAAAA CTCTCAAGGA
5101 TCTTACCGCT GTTGAGATCC AGTTCGATGT AACCCTACTG TGCACCCAAC TGATCTTCAG
5161 CATCTTTTAC TTTCACCAGC GTTTCTGGGT GAGCAAAAAC AGGAAGGCAA AATGCCGCAA
5221 AAAAGGGAAT AAGGGCGACA CGGAAATGTT GAATACTCAT ACTCTTCTT TTTCAATATT
5281 ATTGAAGCAT TTATCAGGGT TATTGTCTCA TGAGCGGATA CATATTGAA TGTATTTAGA
5341 AAAATAAACA AATAGGGGTT CCGCGCACAT TTCCCCGAAA AGTGCCACCT GAAATTGTAA
5401 ACGTTAATAT TTTGTTAAAA TTCGCGTTAA ATTTTGTGTA AATCAGCTCA TTTTTTAAAC
5461 AATAGGCCGA AATCGGCAAA ATCCCTTATA AATCAAAAGA ATAGACCGAG ATAGGGTTGA
5521 GTGTTGTTCC AGTTTGGAAC AAGAGTCCAC TATTAAAGAA CGTGGACTCC AACGTCAAAG
5581 GGGGAAAAAC CGTCTATCAG GCGGATGGCC CACTACGTGA ACCATCACCC TAATCAAGTT
5641 TTTTGGGGTC GAGGTGCCGT AAAGCACTAA ATCGGAACCC TAAAGGGAGC CCCCATTATA
5701 GAGCTTGACG GGGAAAGCCG CGGAACGTGG CGAGAAAGGA AGGGAAGAAA GCGAAAGGAG
5761 CGGGCGCTAG GCGCTGGCG AGTGTAGCGG TCACGCTGCG CGTAACCACC ACACCGCCG
5821 CGCTTAATGC GCGCTACAG GCGCGTCCA TTCGCCATTG AGGCTGCGCA ACTGTTGGGA
5881 AGGGCGATCG GTGCGGCGCT CTTGCTATT ACGCCAGCTG GCGAAAGGGG GATGTGCTGC
5941 AAGGCCGATTA AGTTGGG

FIGURE 26d

Figure 27A: PDEST7

CMV promoter for eukaryotic expression

970 cca ttg acg caa atg ggc ggt agg cgt gta cgg tgg gag gtc tat ata agc
 ggt aac tgc gtt tac ccg cca tcc gca cat gcc acc ctc cag ata tat tgc
 1021 aga gct cgt tta gtg aac ^{→ mRNA start} cgt cag atc gcc tgg aga cgc cat cca cgc tgt
 tct cga gca aat cac ttg gca gtc tag cgg acc tct gcg gta ggt gcg aca
 1072 ttt gac ctc cat aga aga ^{CMV enhancer / promoter} cac cgg gac cga tcc agc ctc cgg act cta gcc
 aaa ctg gag gta tct tct gtg gcc ctg gct agg tgc gag ggc tga gat cgg
 1123 tag gcc gcg gag cgg ata aca att tca cac agg aaa cag cta tga cca cta
 atc cgg cgc ctc gcc tat tgt taa agt gtg tcc ttt gtc gat act ggt gat
 1174 ggc ttt tgc aaa aag cta ttt agg tga cac tat aga agg tac gcc ^{Pst} tgc agg
 ccg aaa acg ttt ttc gat aaa tcc act gtg ata tct tcc atg cgg ^{acg tct}
 1225 ^{Kpn} ^{EcoRI} ^{Int} ^{attR1} ^{Pst}
 tac/cgg tcc gga att ccc atc aca agt tgg ~~tac~~ ~~aaa~~ ~~aaa~~ ggt gaa/cga gaa
 atg gcc agg cct taa ^{ggg} tag tgt tca aac atg ttt ~~tct~~ ~~cga~~ ~~ctc~~ ~~gct~~ ~~ctc~~



pDEST7 6025 bp (rotated to position 2800)

<u>Location (Base Nos.)</u>			<u>Gene Encoded</u>			
	67..589		CMV promoter			
	906..782		attR1			
	1015..1674		CmR			
	1794..1878		inactivated ccdA			
	2016..2321		ccdB			
	2362..2486		attR2			
	2671..3033		small t & polyA			
	3227..3502		f1			
	3962..4822		ampR			
	5022..5661		ori			
1	ATTATCATGA	CATTAACTTA	TAAAAATAGG	CGTAGTACGA	GGCCCTTTCA	CTCATTAGAT
61	GCATGTCGTT	ACATAACTTA	CGGTAAATGG	CCCGCCTGGC	TGACCGCCCA	ACGACCCCCG
121	CCCATTTGACG	TCAATAATGA	CGTATGTTCC	CATAGTAACG	CCAATAGGGA	CTTTCCATTG
181	ACGTCAATGG	GTGGAGTATT	TACGGTAAAC	TGCCCACCTG	GCAGTACATC	AAGTGTATCA
241	TATGCCAAGT	ACGCCCCCTA	TTGACGTCAA	TGACGGTAAA	TGGCCCGCCT	GGCATTATGC
301	CCAGTACATG	ACCTTATGGG	ACTTTCCTAC	TTGGCAGTAC	ATCTACGTAT	TAGTCATCGC
361	TATTACCATG	GTGATGCGGT	TTTGGCAGTA	CATCAATGGG	CGTGGATAGC	GGTTTGACTC
421	ACGGGGATTT	CCAAGTCTCC	ACCCCATTTGA	CGTCAATGGG	AGTTTGTTTT	GGCACCAAAA
481	TCAACGGGAC	TTTCCAAAAT	GTCGTAACAA	CTCCGCCCCA	TTGACGCAAA	TGGGCGGTAG
541	GCGTGACGG	TGGGAGGTCT	ATATAAGCAG	AGCTCGTTTA	GTGAACCGTC	AGATCGCCTG
601	GAGACGCCAT	CCACGCTGTT	TTGACCTCCA	TAGAAGACAC	CGGGACCGAT	CCAGCCTCCG
661	GACTCTAGCC	TAGGCCGCGG	AGCGGATAAC	AATTTACAC	AGGAAACAGC	TATGACCATT
721	AGGCCTTTGC	AAAAAGCTAT	TTAGGTGACA	CTATAGAAGG	TACGCCTGCA	GGTACCGGAT
781	CACAAGTTTG	TACAAAAAAG	CTGAACGAGA	AACGTAAAAT	GATATAAATA	TCAATATATT
841	AAATTAGATT	TTGCATAAAA	AACAGACTAC	ATAATACTGT	AAAACACAAC	ATATCCAGTC
901	ACTATGGCGG	CCGCATTAGG	CACCCAGGC	TTTACACTTT	ATGCTTCCGG	CTCGTATAAT
961	GTGTGGATTT	TGAGTTAGGA	TCCGTCGAGA	TTTTCCAGGAG	CTAAGGAAGC	TAAAAATGGAG
1021	AAAAAATCA	CTGGATATAC	CACCGTTGAT	ATATCCCAAT	GGCATCGTAA	AGAACATTTT
1081	GAGGCATTTT	AGTCAGTTGC	TCAATGTACC	TATAACCAGA	CCGTTTCAGCT	GGATATTACG
1141	GCCTTTTTAA	AGACCGTAAA	GAAAAATAAG	CACAAGTTTT	ATCCGGCCTT	TATTCACATT
1201	CTTGCCCGCC	TGATGAATGC	TCATCCGGAA	TTCCGTATGG	CAATGAAAGA	CGGTGAGCTG
1261	GTGATATGGG	ATAGTGTTCA	CCCTTGTTAC	ACCGTTTTCC	ATGAGCAAAC	TGAAACGTTT
1321	TCATCGCTCT	GGAGTGAATA	CCACGACGAT	TTCCGGCAGT	TTCTACACAT	ATATTCGCAA
1381	GATGTGGCGT	GTTACGGTGA	AAACCTGGCC	TATTTCCCTA	AAGGGTTTTAT	TGAGAATATG
1441	TTTTTCGTCT	CAGCCAATCC	CTGGGTGAGT	TTCAACAGTT	TTGATTTAAA	CGTGGCCAAT
1501	ATGGACAAC	TCTTCGCCCC	CGTTTTTACC	ATGGGCAAAT	ATTATACGCA	AGGCGACAAG
1561	GTGCTGATGC	CGCTGGCGAT	TCAGGTTTCT	CATGCCGTCT	GTGATGGCTT	CCATGTCGGC
1621	AGAATGCTTA	ATGAATTACA	ACAGTACTGC	GATGAGTGGC	AGGGCGGGGC	GTAACCGCGT
1681	GGATCCGGCT	TACTAAAAGC	CAGATAACAG	TATGCGTATT	TGCGCGCTGA	TTTTTGCGGT
1741	ATAAGAATAT	ATACTGATAT	GTATACCCGA	AGTATGTCAA	AAAGAGGTGT	GCTATGAAGC
1801	AGCGTATTAC	AGTGACAGTT	GACAGCGACA	GCTATCAGTT	GCTCAAGGCA	TATATGATGT
1861	CAATATCTCC	GGTCTGGTAA	GCACAACCAT	GCAGAATGAA	GCCCGTCGTC	TGCGTGCCGA
1921	ACGCTGGAAA	GCGGAAAATC	AGGAAGGGAT	GGCTGAGGTC	GCCCGGTTTA	TTGAAATGAA
1981	CGGCTCTTTT	GCTGACGAGA	ACAGGGACTG	GTGAAATGCA	GTTTAAGGTT	TACACCTATA
2041	AAAGAGAGAG	CCGTTATCGT	CTGTTTGTGG	ATGTACAGAG	TGATATTATT	GACACGCCCC
2101	GGCGACGGAT	GGTGTATCCC	CTGGCCAGTG	CACGCTCTGCT	GTCAGATAAA	GTCTCCCGTG
2161	AACCTTACCC	GGTGGTGCAT	ATCGGGGATG	AAAGCTGGCG	CATGATGACC	ACCGATATGG
2221	CCAGTGTGCC	GGTCTCCGTT	ATCGGGGAAG	AAGTGGCTGA	TCTCAGCCAC	CGCGAAAATG
2281	ACATCAAAAA	CGCCATTAAC	CTGATGTTCT	GGGGAATATA	AATGTCAGGC	TCCCTTATAC
2341	ACAGCCAGTC	TGCAGGTGCA	CCATAGTGAC	TGGATATGTT	GTGTTTTACA	GTATTATGTA
2401	GTCTGTTTTT	TATGCAAAAT	CTAATTTAAT	ATATTGATAT	TTATATCATT	TTACGTTTTCT
2461	CGTTCAGCTT	TCTTGTAACA	AGTGGTGATC	GCGTGCATGC	GACGTCATAG	CTCTCTCCCT
2521	ATAGTGAGTC	GTATTATAAG	CTAGGCACTG	GCCGTCGTTT	TACAACGTCG	TGACTGGGAA-

2581 AACTGCTAGC TTGGGATCTT TGTGAAGGAA CCTTACTTCT GTGGTGTGAC ATAATTGGAC
2641 AAACCTACCTA CAGAGATTTA AAGCTCTAAG GTAAATATAA AATTTTTTAAG TGTATAATGT
2701 GTTAAACTAG CTGCATATGC TTGCTGCTTG AGAGTTTTGC TTAGCTAGTA TGATTTATGA
2761 AAATATTATA CACAGGAGCT AGTGATTCTA ATTGTTTGTG TATTTTAGAT TCACAGTCCC
2821 AAGGCTCATT TCAGGCCCTT CAGTCCTCAC AGTCTGTTC A TGATCATAAT CAGCCATACC
2881 ACATTTGTAG AGGTTTTACT TGCTTTAAAA AACCTCCCAC ACCTCCCCCT GAACCTGAAA
2941 CATAAAATGA ATGCAATTGT TGTTGTTAAC TTGTTTATTG CAGCTTATAA TGTTACAAA
3001 TAAAGCAATA GCATCACAAA TTTCACAAAT AAAGCATTTT TTCACTGCA TTCTAGTTGT
3061 GGTTTGTCCA AACTCATCAA TGTATCTTAT CATGTCTGGA TCGATCCTGC ATTAATGAAT
3121 CGGCCAACGC GCGGGGAGAG GCGGTTTTCG TATTGGCTGG CGTAATAGCG AAGAGCCCCG
3181 CACCGATCGC CCTTCCCAAC AGTTGCGCAG CCTGAATGGC GAATGGGACG CGCCCTGTAG
3241 CGGCGCATT A AGCGCGCGG GTGTGGTGGT TACGCGCAGC GTGACCCTA CACTTGCCAG
3301 CGCCCTAGCG CCCGCTCCTT TCGCTTTCTT CCCTTCCTTT CTCGCCAGT TCGCCGGCTT
3361 TCCCCGTCAA GCTCTAAATC GGGGGCTCCC TTTAGGGTTC CGATTTAGTG CTTTACGGCA
3421 CCTCGACCCC AAAAACTTG ATTAGGGTGA TGGTTCACGT AGTGGGCCAT CGCCCTGATA
3481 GACGGTTTTT CGCCCTTTGA CGTTGGAGTC CACGTTCTTT AATAGTGGAC TCTTGTTCGA
3541 AACTGGAACA AACTCAACC CTATCTCGGT CTATCTTTT GATTTATAAG GGATTTTGCC
3601 GATTTCCGCC TATTGGTTAA AAAATGAGCT GATTTAACA AAATTTAAG CGAATTTTAA
3661 CAAAATATTA ACGTTTACAA TTTCAGGTGG CACTTTTCGG GGAATGTGC GCGGAACCCC
3721 TATTTGTTTA TTTTCTAAA TACATTCAAA TATGTATCCG CTCATGCCAG GTCCTGGACT
3781 GGTGAGAACG GCTTGCTCGG CAGCTTCGAT GTGTGCTGGA GGGAGAATAA AGGTCTAAGA
3841 TGTGCGATAG AGGGAAGTCG CATTGAATTA TGTGCTGTGT AGGGATCGCT GGTATCAAT
3901 ATGTGTGCCC ACCCCTGGCA TGAGACAATA ACCCTGATA ATGCTTCAAT AATATTGAAA
3961 AAGGAAGAGT ATGAGTATTC AACATTTCCG TGTCGCCCTT ATTCCCTTTT TTGCGGCATT
4021 TTGCCTTCCT GTTTTGTCTC ACCCAGAAAC GCTGGTGAAA GTAAAAGATG CTGAAGATCA
4081 GTTGGGTGCA CGAGTGGGT ACATCGAACT GGATCTCAAC AGCGGTAAAG TCCTTGAGAG
4141 TTTTCGCCCC GAAGAACGTT TTCCAATGAT GAGCACTTTT AAAGTTCTGC TATGTGGCGC
4201 GGTATTATCC CGTATTGACG CCGGGCAAGA GCAACTCGGT CGCCGCATAC ACTATTCTCA
4261 GAATGACTTG GTTGAGTACT CACCAATCAC AGAAAAGCAT CTTACGGATG GCATGACAGT
4321 AAGAGAATTA TGCAAGTCTG CCATAACCAT GAGTGATAAC ACTGCGGCCA ACTTACTTCT
4381 GACAACGATC GGAGGACCGA AGGAGCTAAC CGCTTTTTTG CACAACATGG GGGATCATGT
4441 AACTCGCCTT GATCGTTGGG AACCGGAGCT GAATGAAGCC ATACCAACG ACGAGCGTGA
4501 CACCAAGATG CCTGTAGCAA TGGCAACAAC GTTGCAGAAA CTATTAAGTG GCGAAGTACT
4561 TACTCTAGCT TCCCGGCAAC AATTAATAGA CTGGATGGAG GCGGATAAAG TTGCAGGACC
4621 ACTTCTGCGC TCGGCCCTTC CGGCTGGCTG GTTTATTGCT GATAAATCTG GAGCCGGTGA
4681 GCGTGGGTCT CGCGGTATCA TTGCAGCACT GGGGCCAGAT GGTAAGCCCT CCCGTATCGT
4741 AGTTATCTAC ACGACGGGGA GTCAGGCAAC TATGGATGAA CGAAATAGAC AGATCGCTGA
4801 GATAGGTGCC TCACTGATTA AGCATTGGTA ACTGTCAGAC CAAGTTTACT CATATATACT
4861 TTAGATTGAT TTAAAACTTC ATTTTAAAT TAAAAGGATC TAGGTGAAGA TCCCTTTTGA
4921 TAATCTCATG CCATAACTTC GTATAATGTA TGCTATACGA AGTTATGGCA TGACCAAAAT
4981 CCCTTAACGT GAGTTTTTCGT TCCACTGAGC GTCAGACCCC GTAGAAAAGA TCAAAGGATC
5041 TTCTTGAGAT CCTTTTTTTC TGCGCGTAAT CTGCTGCTTG CAAACAAAAA AACCACCGCT
5101 ACCAGCGGTG GTTTGTTTGC CGGATCAAGA GCTACCAACT CTTTTTCCGA AGGTAACCTG
5161 CTTACAGAGA GCGCAGATAC CAAATACTGT CCTTCTAGTG TAGCCGTAGT TAGGCCACCA
5221 CTTCAAGAAC TCTGTAGCAC CGCCTACATA CCTCGCTCTG CTAATCCTGT TACCAAGTGGC
5281 TGCTGCCAGT GCGGATAAGT CGTGTCTTAC CGGGTTGGAC TCAAGACGAT AGTTACCGGA
5341 TAAGGCGCAG CGGTCGGGCT GAACGGGGGG TCGTGCACA CAGCCCAGCT TGGAGCGAAC
5401 GACCTACACC GAACTGAGAT ACCTACAGCG TGAGCATTGA GAAAGCGCCA CGCTTCCCGA
5461 AGGAGAAAAG GCGGACAGGT ATCCGGTAAG CGGCAGGGTC GGAACAGGAG AGCGCACGAG
5521 GGAGCTTCCA GGGGGAACG CCTGGTATCT TTATAGTCCT GTCGGGTTTC GCCACCTCTG
5581 ACTTGAGCGT CGATTTTTGT GATGCTCGTC AGGGGGGCGG AGCCTATGGA AAAACGCCAG
5641 CAACGCGGCC TTTTACGGT TCCTGGCCTT TTGCTGGCCT TTTGCTCACA TGTTCTTTCC
5701 TGCGTTATCC CCTGATTCTG TGGATAACCG TATTACCGCC TTTGAGTGAG CTGATACCGC
5761 TCGCCGACG CGAACGACCG AGCGCAGCGA GTCAGTGAGC GAGGAAGCGG AAGAGCGCCC
5821 AATACGCAAA CCGCCTCTCC CCGCGCTTGG GCCGATTAT TAATGCAGAG CTTGCAATTC
5881 GCGCGTTTTT CAATATTATT GAAGCATTTA TCAGGGTTAT TGTCTCATGA GCGGATACAT
5941 ATTTGAATGT ATTTAGAAAA ATAAACAAT AGGGGTTCCG CGCACATTTT CCGGAAAAAT
6001 GCCACCTGAC GTCTAAGAAA CCATT

Figure 78A: pDEST8 Polyhedron Promoter, Baculovirus Transfer Plasmid

AceI

```

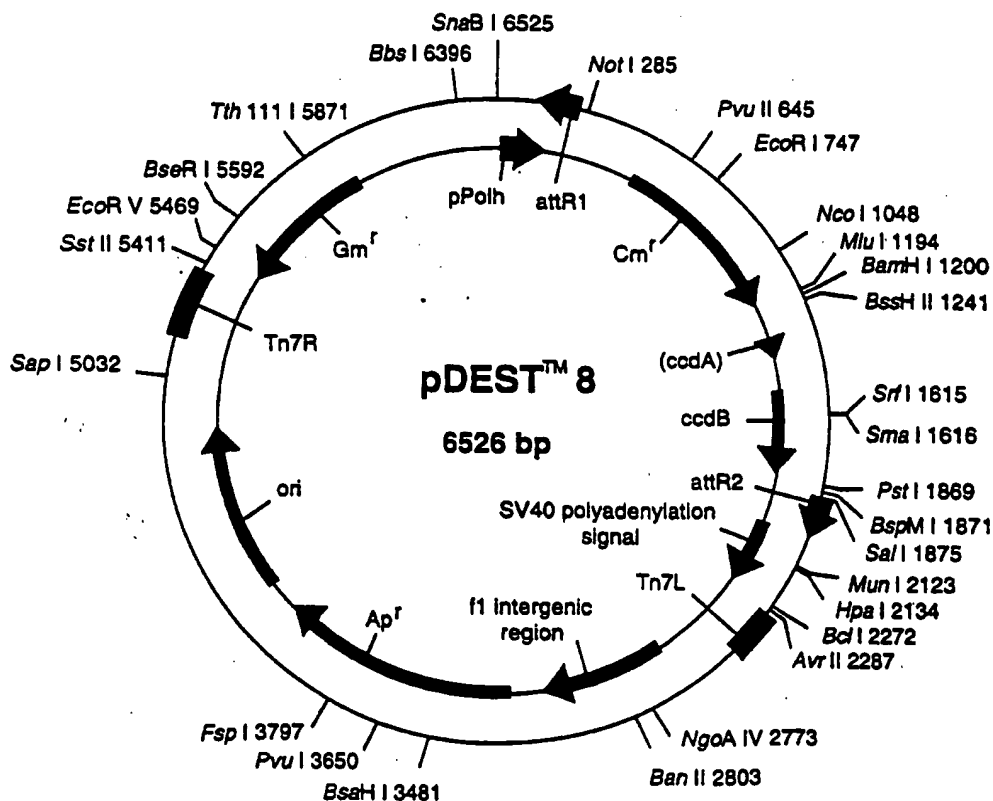
1  cgt|ata ctc cgg aat att aat aga tca tgg aga taa tta aaa tga taa cca
   gca tat gag gcc tta taa tta tct agt acc tct att aat ttt act att ggt
52  tct cgc aaa taa ata agt att tta ctg ttt tct taa cag ttt tgt aat aaa
   aga gcg ttt att tat tca taa aat gac aaa agc att gtc aaa aca tta ttt
103 aaa acc tat aaa tat tcc gga tta ttc ata ccg tcc cac cat cgg gcg cgg
   ttt tgg ata ttt ata agg cct aat aag tat ggc agg gtg gta gcc cgc gcc
154 atc|atc aca agt tgg tac aaa aaa gct gaa cga gaa aag taa tat gat ata
   tag tag tgt tca aac atg ttt ttc cga ctt gct ctt tgc att tta cta tat

```

mRNA (polyhedrin)

attR1

attR2



60/240

pDEST8 6526 bp

<u>Location (Base Nos.)</u>		<u>Gene Encoded</u>
23..152		Ppolh
284..160		attR1
534..1193		CmR
1313..1397		inactivated ccdA
1535..1840		ccdB
1881..2005		attR2
2766..3146		f1
3240..4090		ampR
4289..4869		ori
5564..6496		genR

1	CGTATACTCC	GGAATATTAA	TAGATCATGG	AGATAATTAA	AATGATAACC	ATCTCGCAAA
61	TAAATAAGTA	TTTTACTGTT	TCGTAAACAG	TTTTGTAATA	AAAAAACCTA	TAAATATTCC
121	GGATTATTCA	TACCGTCCCA	CCATCGGGCG	CGGATCATCA	CAAGTTTGTA	CAAAAAAGCT
181	GAACGAGAAA	CGTAAATGA	TATAAATATC	AATATATTAA	ATTAGATTTT	GCATAAAAAA
241	CAGACTACAT	AATACTGTAA	AACACAACAT	ATCCAGTCAC	TATGGCGGCC	GCTAAGTTGG
301	CAGCATCACC	CGACGCACCT	TGCGCCGAAT	AAATACCTGT	GACGGAAGAT	CACTTCGCAG
361	AATAAATAAA	TCCTGGTGTC	CCTGTTGATA	CCGGGAAGCC	CTGGGCCAAC	TTTTGGCGAA
421	AATGAGACGT	TGATCGGCAC	GTAAGAGGTT	CCAACTTTCA	CCATAATGAA	ATAAGATCAC
481	TACCGGGCGT	ATTTTTTGAG	TTATCGAGAT	TTTCAGGAGC	TAAGGAAGCT	AAAATGGAGA
541	AAAAAATCAC	TGGATATACC	ACCGTTGATA	TATCCCAATG	GCATCGTAAA	GAACATTTTG
601	AGGCATTTCA	GTCAGTTGCT	CAATGTACCT	ATAACCAGAC	CGTTCAGCTG	GATATTACGG
661	CCTTTTTTAA	GACCGTAAAG	AAAAATAAGC	ACAAGTTTTA	TCCGGCCTTT	ATTCACATTC
721	TGCCCCGCT	GATGAATGCT	CATCCGGAAT	TCCGTATGGC	AATGAAAGAC	GGTGAGCTGG
781	TGATATGGGA	TAGTGTTCAC	CCTTGTTACA	CCGTTTTCCT	TGAGCAAACT	GAAACGTTTT
841	CATCGCTCTG	GAGTGAATAC	CACGACGATT	TCCGGCAGTT	TCTACACATA	TATTCGCAAG
901	ATGTGGCGTG	TTACGGTGAA	AACCTGGCCT	ATTTCCCTAA	AGGGTTTATT	GAGAATATGT
961	TTTTCGTCTC	AGCCAAATCCC	TGGGTGAGTT	TCACCAAGTTT	TGATTTAAAC	GTGGCCAATA
1021	TGGACAACCT	CTTCGCCCCC	GTTTTACCCA	TGGGCAAAATA	TTATACGCCAA	GGCGACAAGG
1081	TGCTGATGCC	GCTGGCGATT	CAGGTTTCATC	ATGCCGTCTG	TGATGGCTTC	CATGTCGGCA
1141	GAATGCTTAA	TGAATTACAA	CAGTACTGCG	ATGAGTGGCA	GGGCGGGGCG	TAAACCGCTG
1201	GATCCGGCTT	ACTAAAAGCC	AGATAACAGT	ATGCGTATTT	GCGCGCTGAT	TTTTGCGGTA
1261	TAAGAATATA	TACTGATATG	TATACCCGAA	GTATGTCAAA	AAGAGGTGTG	CTATGAAGCA
1321	GCGTATTACA	GTGACAGTTG	ACAGCGACAG	CTATCAGTTG	CTCAAGGCAT	ATATGATGTC
1381	AATATCTCCG	GTCTGGTAAG	CACAACCATG	CAGAATGAAG	CCCGTCGTCT	GCGTGCCGAA
1441	CGCTGGAAG	CGGAAAATCA	GGAAGGGATG	GCTGAGGTG	CCCGGTTTAT	TGAAATGAAC
1501	GGCTCTTTTG	CTGACGAGAA	CAGGGACTGG	TGAAATGCAG	TTTAAGGTTT	ACACCTATAA
1561	AAGAGAGAGC	CGTTATCGTC	TGTTTGTGGA	TGTACAGAGT	GATATTATTG	ACACGCCCGG
1621	GCGACGGATG	GTGATCCCCC	TGGCCAGTGC	ACGTCTGCTG	TCAGATAAAG	TCTCCCGTGA
1681	ACTTTACCCG	GTGGTGATA	TCGGGGATGA	AAGCTGGCGC	ATGATGACCA	CCGATATGGC
1741	CAGTGTGCCG	GTCTCCGTTA	TCGGGGAAGA	AGTGGCTGAT	CTCAGCCACC	GCGAAAATGA
1801	CATCAAAAAC	GCCATTAACC	TGATGTTCTG	GGGAATATAA	ATGTCAGGCT	CCCTTATACA
1861	CAGCCAGTCT	GCAGGTCGAC	CATAGTGACT	GGATATGTTG	TGTTTTACAG	TATTATGTAG
1921	TCTGTTTTTT	ATGCAAAATC	TAATTTAATA	TATTGATATT	TATATCATT	TACGTTTCTC
1981	GTTTACGCTT	CTTGTAACAA	GTGGTGATAG	CTTGTCGAGA	AGTACTAGAG	GATCATAATC
2041	AGCCATACCA	CATTTGTAGA	GGTTTACTT	GCTTTAAAAA	ACCTCCACA	CCTCCCCCTG
2101	AACCTGAAAC	ATAAAATGAA	TGCAATTGTT	GTTGTTAACT	TGTTTATTGC	AGCTTATAAT
2161	GGTTACAAAT	AAAGCAATAG	CATCACAAT	TTCAACAATA	AAGCATTTTT	TTCATGTCAT
2221	TCTAGTTGTG	GTTTGTCCAA	ACTCATCAAT	GTATCTTATC	ATGCTGGAT	CTGATCACTG
2281	CTTGAGCCTA	GGAGATCCGA	ACCAGATAAG	TGAAATCTAG	TTCCAAACTA	TTTTGTCAAT
2341	TTTAATTTTC	GTATTAGCTT	ACGACGCTAC	ACCCAGTTCC	CATCTATTTT	GTCACCTCTC
2401	CCTAAATAAT	CCTTAAAAAC	TCCATTTCCA	CCCCTCCAG	TTCCCAACTA	TTTTGTCCGC
2461	CCACAGCGGG	GCATTTTCT	TCCTGTTATG	TTTTTAATCA	AACATCCTGC	CAACTCCATG
2521	TGACAAACCG	TCATCTTCGG	CTACTTTTTT	TCTGTCACAG	AATGAAAATT	TTTCTGTCAT

FIGURE 28B

61/240

2581 CTCTTCGTTA TTAATGTTTG TAATTGACTG AATATCAACG CTTATTTGCA GCCTGAATGG
2641 CGAATGGACG CGCCCTGTAG CGGCGCATTA AGCGCGGCGG GTGTGGTGGT TACGCGCAGC
2701 GTGACCGCTA CACTTGCCAG CGCCCTAGCG CCCGCTCCTT TCGCTTTCTT CCCTTCCTTT
2761 CTCGCCACGT TCGCCGGCTT TCCCCGTCAG GCTCTAAATC GGGGGCTCCC TTTAGGGTTC
2821 CGATTTAGTG CTTTACGGCA CCTCGACCCC AAAAACTTG ATTAGGGTGA TGGTTCACGT
2881 AGTGGGCCAT CGCCCTGATA GACGGTTTTT CGCCCTTGA CGTTGGAGTC CACGTCTTTT
2941 AATAGTGGAC TCTTGTTCCA AACTGGAACA AACTCAACC CTATCTCGGT CTATTCTTTT
3001 GATTTATAAG GGATTTTGCC GATTTTCGGC TATTGGTTAA AAAATGAGCT GATTTAACAA
3061 AAATTTAACG CGAATTTTAA CAAATATTA ACGTTTACAA TTTCAGGTGG CACTTTTCGG
3121 GGAAATGTGC GCGGAACCCC TATTTGTTTA TTTTCTAAA TACATTCAA TATGTATCCG
3181 CTCATGAGAC AATAACCTG ATAAATGCTT CAATAATATT GAAAAAGGAA GAGTATGAGT
3241 ATTCAACATT TCCGTGTCGC CCTTATTCCT TTTTTCGGG CATTTCGCCT TCCTGTTTTT
3301 GCTCACCCAG AAACGCTGGT GAAAGTAAAA GATGCTGAAG ATCAGTTGGG TGCACGAGTG
3361 GGTTACATCG AACTGGATCT CAACAGCGGT AAGATCCTTG AGAGTTTTCG CCCCAGAGAA
3421 CGTTTTCCAA TGATGAGCAC TTTTAAAGTT CTGCTATGTG GCGCGGTATT ATCCCGTATT
3481 GACGCCGGGC AAGAGCAACT CGGTGCGCGC ATACACTATT CTCAGAATGA CTTGGTTGAG
3541 TACTCACCAG TCACAGAAAA GCATCTTACG GATGGCATGA CAGTAAGAGA ATTATGCAGT
3601 GCTGCCATAA CCATGAGTGA TAACACTGCG GCCAATTAC TTCTGACAAC GATCGGAGGA
3661 CCGAAGGAGC TAACCGCTTT TTTGCACAAC ATGGGGGATC ATGTAACCTG CCTTGATCGT
3721 TGGGAACCGG AGCTGAATGA AGCCATACCA AACGACGAGC GTGACACCAC GATGCCTGTA
3781 GCAATGGCAA CAACGTTGCG CAACTATTA ACTGGCGAAC TACTTACTCT AGCTTCCCGG
3841 CAACAATTAA TAGACTGGAT GGAGGCGGAT AAAGTTGCAG GACCCTTCT GCGCTCGGCC
3901 CTTCCGGCTG GCTGGTTTAT TGCTGATAAA TCTGGAGCGG GTGAGCGTGG GTCTCGCGGT
3961 ATCATTGTCAG CACTGGGGCC AGATGGTAAG CCTCCCGTA TCGTAGTTAT CTACACGACG
4021 GGGAGTCAGG CAACTATGGA TGAACGAAAT AGACAGATCG CTGAGATAGG TGCCTCACTG
4081 ATTAAGCATT GGTAAGTGT AGACCAAGTT TACTCATATA TACTTTAGAT TGATTTAAAA
4141 CTTTATTTTT AATTTAAAAG GATCTAGGTG AAGATCCTTT TTGATAATCT CATGACCAAA
4201 ATCCCTTAAC GTGAGTTTTT GTTCCACTGA CCGTCAGACC CCGTAGAAAA GATCAAAAGGA
4261 TCTTCTTGAG ATCCTTTTTT TCTGCGGTA ATCTGCTGCT TGCAACAAAA AAAACCACCG
4321 CTACCAGCGG TGGTTTGTG GCCGGATCAA GAGCTACCAA CTCTTTTTCC GAAGGTAAGT
4381 GGCTTCAGCA GAGCGCAGAT ACCAAATACT GTCCTTCTAG TGTAGCCGTA GTTAGGCCAC
4441 CACTTCAAGA ACTCTGTAGC ACCGCTACA TACCTCGCTC TGCTAATCCT GTTACCAGTG
4501 GCTGCTGCCA GTGGCGATAA GTCGTGTCTT ACCGGGTTGG ACTCAAGACG ATAGTTACCG
4561 GATAAGGCGC AGCGGTGCGG CTGAACGGGG GGTTCGTGCA CACAGCCGAG CTTGGAGCGA
4621 ACGACCTACA CCGAACTGAG ATACCTACAG CGTGAGCATT GAGAAAGCGC CACGCTTCCC
4681 GAAGGGAGAA AGGCGGACAG GTATCCGGTA AGCGGCAGGG TCGGAACAGG AGAGCGCACG
4741 AGGGAGCTTC CAGGGGGAAA CGCCTGGTAT CTTTATAGTC CTGTGCGGTT TCGCCACCTC
4801 TGACTTGAGC GTCGATTTTT GTGATGCTCG TCAGGGGGGG GGAGCCTATG GAAAAACGCC
4861 AGCAACGCGG CTTTTCACG GTTCTTGCGC TTTTGCTGGC CTTTGTCTCA CATGTTCTTT
4921 CCTGCGTTAT CCCCTGATTC TGTGGATAAC CGTATTACCG CCTTTGAGTG AGCTGATACC
4981 GCTCGCCGCA GCCGAACGAC CGAGCGCAGC GAGTCAGTGA GCGAGGAAGC GGAAGAGCGC
5041 CTGATGCGGT ATTTTCTCCT TACGCATCTG TGCGGTATTT CACACCGCAG ACCAGCCGCG
5101 TAACCTGGCA AAATCGGTTA CGGTTGAGTA ATAAATGGAT GCCCTGCGTA AGCGGCTGTG
5161 GGCGGACAAT AAAGTCTTAA ACTGAACAAA ATAGATCTAA ACTATGACAA TAAAGTCTTA
5221 AACTAGACAG AATAGTTGTA AACTGAAATC AGTCCAGTTA TGCTGTGAAA AAGCATACTG
5281 GACTTTTGTT ATGGCTAAAG CAACTCTTC ATTTTCTGAA GTGCAAATTG CCCGTGCTAT
5341 TAAAGAGGGG CGTGGCCAAG GGCATGGTAA AGACTATATT CGCGGCGTTG TGACAATTTA
5401 CCGAACAAC CCGCGGCCGG GAAGCCGATC TCGGCTTGAA CGAATTGTTA GGTGGCGGTA
5461 CTTGGGTCGA TATCAAGTG CATCACTTCT TCCCGTATGC CCAACTTTGT ATAGAGAGCC
5521 ACTGCGGGAT CGTCACCGTA ATCTGCTTGC ACGTAGATCA CATAAGCACC AAGCGCGTTG
5581 GCCTCATGCT TGAGGAGATT GATGAGCGCG GTGGCAATGC CCTGCCTCCG GTGCTCGCCG
5641 GAGACTGCGA GATCATAGAT ATAGATCTCA CTACGCGGCT GCTCAAACCT GGGCAGAACG
5701 TAAGCCGCGA GAGCGCCAAC AACCGCTTCT TGGTCGAAGG CAGCAAGCGC GATGAATGTC
5761 TTAAGTACGGA GCAAGTTCCC GAGGTAATCG GAGTCCGGCT GATGTTGGGA GTAGGTGGCT
5821 ACGTCTCCGA ACTCACGACC GAAAAGATCA AGAGCAGCCC GCATGGATTT GACTTGGTCA
5881 GGGCCGAGCC TACATGTGCG AATGATGCCC ATACTTGAGC CACCTAACTT TGTTTTAGGG
5941 CGACTGCCCT GCTGCGTAAC ATCGTTGCTG CTGCGTAACA TCGTTGCTGC TCCATAACAT
6001 CAAACATCGA CCCACGGCGT AACGCGCTTG CTGCTTGGAT GCCCGAGGCA TAGACTGTAC-

FIGURE 28C

62/240

6061 AAAAAACAG TCATAACAAG CCATGAAAAC CGCCACTGCG CCGTTACCAC CGCTGCGTTC
6121 GGTCAAGGTT CTGGACCAGT TGGGTGAGCG CATACGCTAC TTGCATTACA GTTTACGAAC
6181 CGAACAGGCT TATGTCAACT GGGTTCGTGC CTTCATCCGT TTCCACGGTG TCGGTCACCC
6241 GGCAACCTTG GGCAGCAGCG AAGTCGAGGC ATTTCTGTCC TGGCTGGCGA ACGAGCGCAA
6301 GGTTTCGGTC TCCACGCATC GTCAGGCATT GGCGGCCTTG CTGTTCTTCT ACGGCAAGGT
6361 GCTGTGCACG GATCTGCCCT GGCTTCAGGA GATCGGAAGA CCTCGGCCGT CGCGGCGCTT
6421 GCCGGTGGTG CTGACCCCGG ATGAAGTGGT TCGCATCCTC GGTTTCTGG AAGGCGAGCA
6481 TCGTTTGTTT GCCCAGGACT CTAGCTATAG TTCTAGTGGT TGGCTA

FIGURE 28D

63/240

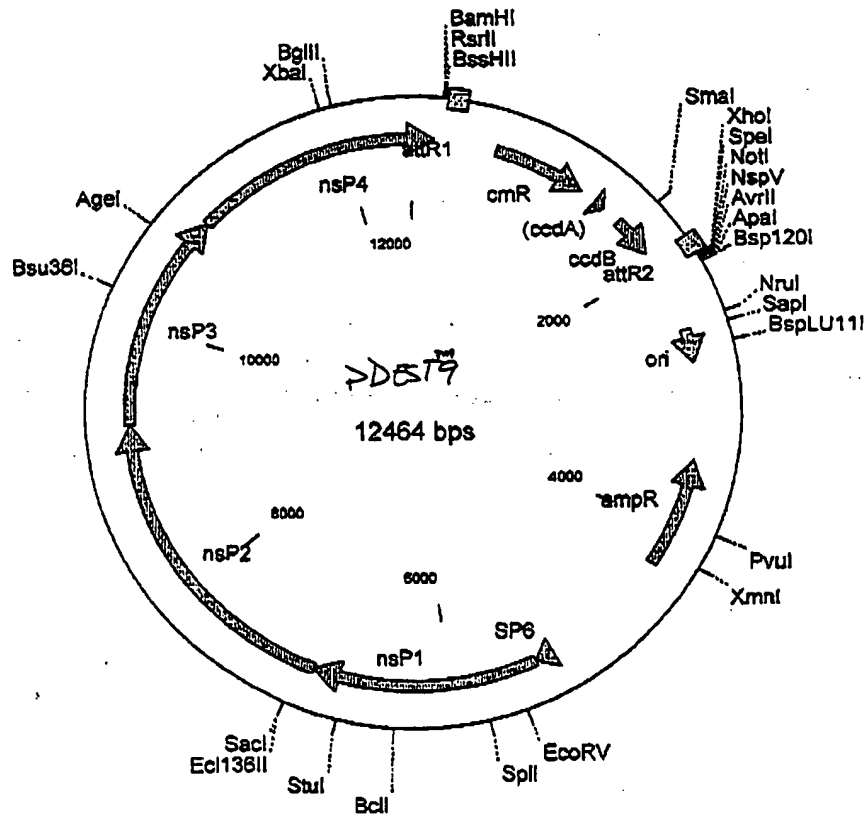
Figure 29A: pDEST9

Semliki Forest Virus vector

103 ttg gcg agg gac att aag gcg ttt aag aaa ttg aga gga cct gtt ata ~~cac~~
 aac cgc tcc ctg taa ttc cgc aaa ttc ttt aac tct cct gga caa tat ~~gtg~~
 243 promoter → 245 RNA

154 ~~ctc tac ggc ggt cct aga ttg gtc cgt taa tac aca gaa ttc tga ttg gat~~ Bam
~~gag atg ccg cca gga tct aac cac gca att atg tgt ctt aag act aac cta~~

205 ~~ccc ggt ccg aag cgc gct ttc cca tca~~ ~~aca agt tct tac aac aac gct gga~~ ~~att R1~~
~~ggg cca ggc ttc gcg cga aag ggt agt~~ ~~tgt tca aac atg ttt tta cga cta~~



64/240

pDEST9 12464 bp

<u>Location (Base Nos.)</u>		<u>Gene Encoded</u>
355..232		attR1
605..1264		CmR
1384..1468		inactivated ccdA
1606..1911		ccdB
1952..2078		attR2
2532..2782		ori
3482..4282		ampR
5232..5365		SP6 promoter
5365..6965		nsP1:non-structural protein 1
6965..9265		nsP2:non-structural protein 2
9265..10865		nsP3:non-structural protein 3
10865..161		nsP4:non-structural protein 4

1	AGCAAGTGGT	TCCGGACAGG	CTTGGGGGCC	GAAGTGGAGG	TGGCACTAAC	ATCTAGGTAT
61	GAGGTAGAGG	GCTGCAAAAG	TATCCTCATA	GCCATGGCCA	CCTTGGCGAG	GGACATTAAG
121	GCGTTTAAGA	AATTGAGAGG	ACCTGTTATA	CACCTCTACG	GCGGTCCTAG	ATTGGTGCCT
181	TAATACACAG	AATTCTGATT	GGATCCCGGT	CGAAGCGCG	CTTTCCCATC	ACAAGTTTGT
241	ACAAAAAGC	TGAACGAGAA	ACGTAAAATG	ATATAAATAT	CAATATATTA	AATTAGATTT
301	TGCATAAAAA	ACAGACTACA	TAATACTGTA	AAACACAACA	TATCCAGTCA	CTATGGCGGC
361	CGCTAAGTTG	GCAGCATCAC	CCGACGCACT	TTGCGCCGAA	TAAATACCTG	TGACGGAAGA
421	TCACCTCGCA	GAATAAATA	ATCCTGGTGT	CCCTGTTGAT	ACCGGGAAGC	CCTGGGCCAA
481	CTTTTGGCGA	AAATGAGACG	TTGATCGGCA	CGTAAGAGGT	TCCAACCTTC	ACCATAATGA
541	AATAAGATCA	CTACCGGGCG	TATTTTTTGA	GTTATCGAGA	TTTTTCAGGAG	CTAAGGAAGC
601	TAAATGGAG	AAAAAATCA	CTGGATATAC	CACCGTTGAT	ATATCCCAAT	GGCATCGTAA
661	AGAACATTTT	GAGGCATTTC	AGTCAGTTGC	TCAATGTACC	TATAACCAGA	CCGTTTCAGCT
721	GGATATTACG	GCCTTTTTTA	AGACCGTAAA	GAAAAATAAG	CACAAGTTTT	ATCCGGCCTT
781	TATTCACATT	CTTGCCCGCC	TGATGAATGC	TCATCCGGAA	TTCCGTATGG	CAATGAAAGA
841	CGGTGAGCTG	GTGATATGGG	ATAGTGTTC	CCCTTGTTAC	ACCGTTTTCC	ATGAGCAAAC
901	TGAAACGTTT	TCATCGCTCT	GGAGTGAATA	CCACGACGAT	TTCCGGCAGT	TTCTACACAT
961	ATATTCGCAA	GATGTGGCGT	GTTACGGTGA	AAACCTGGCC	TATTTCCCTA	AAGGGTTTAT
1021	TGAGAATATG	TTTTTCGTCT	CAGCCAATCC	CTGGGTGAGT	TTCACCAGTT	TTGATTTAAA
1081	CGTGGCCAAAT	ATGGACAAC	TCTTCGCCCC	CGTTTTTACC	ATGGGCAAAT	ATTATACGCA
1141	AGGCGACAAG	GTGCTGATGC	CGCTGGCGAT	TCAGGTTTCAT	CATGCCGTCT	GTGATGGCTT
1201	CCATGTCGGC	AGAATGCTTA	ATGAATTACA	ACAGTACTGC	GATGAGTGGC	AGGGCGGGGC
1261	GTAAGATCT	GGATCCGGCT	TACTAAAAGC	CAGATAACAG	TATGCGTATT	TGCGCGTGTA
1321	TTTTTGCGGT	ATAAGATAT	ATACTGATAT	GTATACCCGA	AGTATGTCAA	AAAGAGGTGT
1381	GCTATGAAGC	AGCGTATTAC	AGTGACAGTT	GACAGCGACA	GCTATCAGTT	GCTCAAGGCA
1441	TATATGATGT	CAATATCTCC	GGTCTGGTAA	GCACAACCAT	GCAGAATGAA	GCCCCTCGTC
1501	TGCGTGCCGA	ACGCTGGAAA	GCGGAAAATC	AGGAAGGGAT	GGCTGAGGTC	GCCCGGTTTA
1561	TTGAAATGAA	CGGCTCTTTT	GCTGACGAGA	ACAGGGACTG	GTGAAATGCA	GTTTAAGGTT
1621	TACACCTATA	AAAGAGAGAG	CCGTTATCGT	CTGTTTGTGG	ATGTACAGAG	TGATATTATT
1681	GACACGCCCG	GGCGACGGAT	GGTGATCCCC	CTGGCCAGTG	CACGTCTGCT	GTGAGATAAA
1741	GTCTCCCGTG	AACTTTACCC	GGTGGTGCAT	ATCGGGGATG	AAAGCTGGCG	CATGATGACC
1801	ACCGATATGG	CCAGTGTGCC	GGTCTCCGTT	ATCGGGGAAG	AAGTGGCTGA	TCTCAGCCAC
1861	CGCGAAAATG	ACATCAAAAA	CGCCATTAAAC	CTGATGTTCT	GGGGAATATA	AATGTCAGGC
1921	TCCCTTATAC	ACAGCCAGTC	TGCAGGTGCA	CCATAGTGAC	TGGATATGTT	GTGTTTTACA
1981	GTATTATGTA	GTCTGTTTTT	TATGCAAAAG	TGCTAATTTA	ATATATTGAT	ATTTATATCA
2041	TTTTACGTTT	CTCGTTCAGC	TTTCTTGTTAC	AAAGTGGTGA	TGGGAACCTG	AGTTCAC TAG
2101	TCGATCCCGC	GGCCGCTTTC	GAACCTAGGC	AAGCATGCGG	GCCCAGTGGG	TAATTAATTG
2161	AATTACATCC	CTACGCAAAAC	GTTTTACGGC	CGCCGCTGGC	GCCCAGCGCC	GGCCGCCCCG
2221	CCTTGGCCGT	TGCAGGCCAC	TCCGGTGGCT	CCCGTCGTCC	CCGACTTCCA	GGCCAGCAG
2281	ATGCAGCAAC	TCATCAGCGC	CGTAAATGCG	CTGACAATGA	GACAGAACGC	AATTGCTCCT
2341	GCTAGGAGCT	TAATTCGACG	AATAATTGGA	TTTTTATTTT	ATTTTGCAAT	TGGTTTTTAA
2401	TATTTCCAAA	AAAAAATAAA	AAAAAATAAA	AAAAAATAAA	AAAAAATAAA	AAAAAATAAA

FIGURE 29B

2461 AAAAAAAAAA AAAAAAATA GAAATCGCGA TTTCTAGTCT GCATTAATGA ATCGGCCAAC
2521 GCGCGGGGAG AGGCGGTTTG CGTATTGGGC GCTCTTCCGC TTCCTCGCTC ACTGACTCGC
2581 TGCGCTCGGT CGTTCCGGCTG CGGCGAGCGG TATCAGCTCA CTCAAAGGCG GTAATACGGT
2641 TATCCACAGA ATCAGGGGAT AACGCAGGAA AGAACATGTG AGCAAAAGGC CAGCAAAAGG
2701 CCAGGAACCG TAAAAAGGCC GCGTTGCTGG CGTTTTTCCA TAGGCTCCGC CCCCTGACG
2761 AGCATCACAA AAATCGACGC TCAAGTCAGA GGTGGCGAAA CCCGACAGGA CTATAAAGAT
2821 ACCAGGCGTT TCCCCCTGGA AGCTCCCTCG TCGCTCTCC TGTTCCGACC CTGCCGCTTA
2881 CCGGATACCT GTCCGCCTTT CTCCCTTCGG GAAGCGTGGC GCTTTCTCAA TGCTCGCGCT
2941 GTAGGTATCT CAGTTCGGTG TAGGTCGTTT GCTCCAAGCT GGGCTGTGTG CACGAACCCC
3001 CCGTTCAGCC CGACCGCTGC GCCTTATCCG GTAACATATCG TCTTGAGTCC AACCCGGTAA
3061 GACACGACTT ATCGCCACTG GCAGCAGCCA CTGGTAACAG GATTAGCAGA GCGAGGTATG
3121 TAGGCGGTGC TACAGAGTTC TTGAAGTGGT GGCCTAACTA CGGCTACACT AGAAGGACAG
3181 TATTTGGTAT CTGCGCTCTG CTGAAGCCAG TTACCTTCGG AAAAAGAGTT GGTAGCTCTT
3241 GATCCGGCAA ACAAACCACC GCTGGTAGCG GTGGTTTTTT TGTTTGCAAG CAGCAGATTA
3301 CCGCAGAAA AAAAGGATCT CAAGAAGATC CTTTGATCTT TTCTACGGGG TCGTCAGCTC
3361 AGTGAACGA AAACCTCACG TAAGGGATTT TGGTCATGAG ATTATCAAAA AGGATCTTCA
3421 CCTAGATCCT TTTAAATTAA AAATGAAGTT TTAATCAAT CTAAAGTATA TATGAGTAAA
3481 CTTGGTCTGA CAGTTACCAA TGCTTAATCA GTGAGGCACC TATCTCAGCG ATCTGTCTAT
3541 TTCGTTTATC CATAGTTGCC TGAATCCCG TCGTGTAGAT AACTACGATA CGGGAGGGCT
3601 TACCATCTGG CCCAGTGCT GCAATGATAC CGCGAGACCC ACGCTCACCG GCTCCAGATT
3661 TATCAGCAAT AAACCAGCCA GCCGGAAGGG CCGAGCGCAG AAGTGGTCTT GCAACTTTAT
3721 CCGCCTCCAT CCAGTCTATT AATTGTTGCC GGGAAAGCTAG AGTAAAGTAGT TCGCAGTTA
3781 ATAGTTTTCG CAACGTTGTT GCCATTGCTA CAGGCATCGT GGTGTACGC TCCTCGTTTG
3841 GTATGGCTTC ATTCAGCTCC GGTTCCTAAC GATCAAGGCG AGTTACATGA TCCCCATGT
3901 TGTGCAAAAA AGCGGTTAGC TCCTTCGGTC CTCCGATCGT TGTCAGAAGT AAGTTGGCCG
3961 CAGTGTATAT ACTCATGTT ATGGCAGCAC TGCATAATTC TCTTACTGTC ATGCCATCCG
4021 TAAGATGCTT TTCTGTGACT GGTGAGTACT CAACCAAGTC ATTCTGAGAA TAGTGTATGC
4081 GGGCACCAGG TTGCTCTTGC CCGGCGTCAA TACGGGATAA TACCGGCCA CATAGCAGAA
4141 CTTTAAAAGT GCTCATATT GAAGAAAGTT CTTGCGGGCG AAAACTCTCA AGGATCTTAC
4201 CGCTGTTGAG ATCCAGTTCG ATGTAACCCA CTCGTGCACC CAACTGATCT TCAGCATCTT
4261 TTACTTTTAC CAGCGTTTCT GGTGAGCAA AAACAGGAAG GCAAAATGCC GCAAAAAAGG
4321 GAATAAGGGC GACACGAAA TGTTGAATAC TCATACTCTT CCTTTTCAA TATTATTGAA
4381 GCATTTATCA GGGTTATTGT CTGATGAGCG GATACATATT TGAATGTATT TAGAAAAATA
4441 AACAAATAGG GGTTCCGCGC ACATTTCCCC GAAAAGTGCC ACCTGACGTC TAAGAAACCA
4501 TTATTATCAT GACATTAACC TATAAAAATA GCGGTATCAC GAGGCCCTTT CGTCTCGCGC
4561 GTTTCGGTGA TGACGGTGAA AACCTCTGAC ACATGCAGCT CCCGAGACG GTCACAGCTT
4621 CTGTCTAAGC GGATGCCGGG AGCAGACAAG CCCGTCAGGG CGCGTCAGCG GGTGTTGGCG
4681 GGTGTCGGGG CTGGCTTAAC TATGCGGCAT CAGAGCAGAT TGTACTGAGA GTGCACCATA
4741 TCGACGCTCT CCCTTATGCG ACTCCTGCAT TAGGAAGCAG CCCAGTACTA GGTGAGGCC
4801 GTTGAGCACC GCCGCCGCAA GGAATGGTGC ATGCAAGGAG ATGGCGCCCA ACAGTCCCCC
4861 GGCCACGGGG CCTGCCACCA TACCCACGCC GAAACAAGCG CTCATGAGCC CGAAGTGGCG
4921 AGCCCGATCT TCCCCATCGG TGATGTCGGC GATATAGGCG CCAGCAACCG CACCTGTGGC
4981 GCCGGTGATG CCGGCCACGA TCGGTCCGGC GTAGAGGATC TGGCTAGCGA TGACCCTGCT
5041 GATTGGTTTC CTGACCATTT CCGGGGTGCG GAACGGCGTT ACCAGAAACT CAGAAGGTTT
5101 GTCCAACCAA ACCGACTCTG ACGGCAGTTT ACAGAGAGA TGATAGGGTC TGCTTCAGTA
5161 AGCCAGATGC TACACAATTA GGCTTGTA CA TATTGTCGTT AGAACCGCGC TACAATTAAT
5221 ACATAACCTT ATGTATCATA CACATACGAT TTAGGTGACA CTATAGATGG CGGATGTGTG
5281 ACATACACGA CGCCAAAAGA TTTGTTCCA GCTCCTGCCA CCTCCGCTAC GCGAGAGATT
5341 AACCACCCAC GATGGCCGCC AAAGTGCATG TTGATATTGA GGCTGACAGC CCATTCATCA
5401 AGTCTTTGCA GAAGGCATTT CCGTCGTTCC AGGTGGAGTC ATTGCAGGTC ACACCAAATG
5461 ACCATGCAAA TGCCAGAGCA TTTTCGACC TGGCTACCAA ATTGATCGAG CAGGAGACTG
5521 ACAAAGACAC ACTCATCTTG GATATCGGCA GTGCGCCTTC CAGGAGAATG ATGTCTACGC
5581 ACAAATACCA CTGCGTATGC CCTATGCGCA GCGCAGAAGA CCCCAGAAAG CTCGATAGCT
5641 ACGCAAAGAA ACTGGCAGCG GCCTCCGGGA AGGTGCTGGA TAGAGAGATC GCAGGAAAAA
5701 TCACCGACCT GCAGACCGTC ATGGCTACGC CAGACGCTGA ATCTCCTACC TTTTGCTGCTG
5761 ATACAGACGT CACGTGTCGT ACGGCAGCCG AAGTGGCCGT ATACCAGGAC GTGTATGCTG
5821 TACATGCACC AACATCGCTG TACCATCAGG CGATGAAAGG TGTGAAAGC GCGTATTGGA
5881 TTGGGTTTGA CACCACCCCG TTTATGTTTG ACGCGCTAGC AGGCGCGTAT CCAACCTACG-

66/240

5941 CCACAAACTG GGCCGACGAG CAGGTGTTAC AGGCCAGGAA CATAGGACTG TGTGCASCAT
 6001 CCTTGACTGA GGAAGACTC GGCAAACTGT CCATTCTCCG CAAGAAGCAA TTGAAACCTT
 6061 GCGACACAGT CATGTTCTCG GTAGGATCTA CATTGTACAC TGAGAGCAGA AAGCTACTGA
 6121 GGAGCTGGCA CTTACCCTCC GTATTCCACC TGAAAGGTAA ACAATCCTTT ACCTGTAGGT
 6181 GCGATACCAT CGTATCATGT GAAGGGTACG TAGTTAAGAA AATCACTATG TGCCCCGGCC
 6241 TGTACGGTAA AACGGTAGGG TACGCCGTGA CGTATCACGC GGAGGGATTG CTAGTGTGCA
 6301 AGACCACAGA CACTGTCAA GGAGAAAGAG TCTCATTCCC TGTATGCACC TACGTCCCCT
 6361 CAACCATCTG TGATCAAAATG ACTGGCATAC TAGCGACCGA CGTCACACCG GAGGACGCAC
 6421 AGAAGTTGTT AGTGGGATTG AATCAGAGGA TAGTTGTGAA CGGAAGAACA CAGCGAAACA
 6481 CTAACACGAT GAAGAACTAT CTGCTTCCGA TTGTGGCCGT CGCATTTAGC AAGTGGSCGA
 6541 GGAATACAA GGCAGACCTT GATGATGAAA AACCTCTGGG TGTCCGAGAG AGGTCACTTA
 6601 CTGTCTGCTG CTTGTGGGCA TTTAAACGA GGAAGATGCA CACCATGTAC AAGAAACAG
 6661 ACACCCAGAC AATAGTGAAG GTGCCCTCAG AGTTTAACTC GTTCGTATC CCGAGCCTAT
 6721 GGTCTACAGG CCTCGCAATC CCAGTCAGAT CACGCATTAA GATGCTTTTG GCCAAGAAGA
 6781 CCAAGCGAGA GTTAATACCT GTTCTCGAGC CGTCGTCAGC CAGGGATGCT GAACAAGSAGG
 6841 AGAAGGAGAG GTTGGAGGCC AAGCTGACTA GAGAAGCCTT ACCACCCCTC GTCCCAATCG
 6901 CGCCGGCGGA GACGGGAGTC GTCGACGTCG ACGTTGAAGA ACTAGAGTAT CACGCAAGTG
 6961 CAGGGGTCTG GGAACACCT CGCAGCGCGT TGAAAGTCAC CGCACAGCCG AACGACSTAC
 7021 TACTAGGAAA TTACGTAGTT CTGTCCCCGC AGACCGTGCT CAAGAGCTCC AAGTTGSCCC
 7081 CCGTGACCCC TCTAGCAGAG CAGGTGAAAA TAATAACACA TAACGGGAGG GCCGGCGGTT
 7141 ACCAGGTCGA CGGATATGAC GGCAGGGTCC TACTACCATG TGGATCGGCC ATTCGGSTCC
 7201 CTGAGTTTCA GGCTTTGAGC GAGAGCGCCA CTATGGTGTA CAACGAAAGG GAGTTCTCTA
 7261 ACAGGAAACT ATACCATATT GCCGTTACG GACCTCGCT GAACACCGAC GAGGAGAACT
 7321 ACGAGAAAGT CAGAGCTGAA AGAACTGACG CCGAGTACGT GTTCGACGTA GATAAAAAAT
 7381 GCTGCGTCAA GAGAGAGGAA GCGTCGGGTT TGGTGTGGT GGGAGAGCTA ACCAACCCCC
 7441 CGTTCATGA ATTGCGCTAC GAAGGGCTGA AGATCAGGCC GTCGGCACA TATAAGACTA
 7501 CAGTAGTAGG AGTCTTTGGG GTTCCGGGAT CAGGCAAGTC TGCTATTATT AAGAGCCTCG
 7561 TGACCAACA CGATCTGGTC ACCAGCGGCA AGAAGGAGAA CTGCCAGGAA ATAGTTAAGC
 7621 ACGTGAAGAA GCACCGCGGG AAGGGGACAA GTAGGGAAAA CAGTGACTCC ATCTGTCTAA
 7681 ACGGGTGTG TCGTGCCGTG GACATCTAT ATGTGGACGA GGCTTTCTGCT TGCCATTCCG
 7741 GTACTCTGCT GGCCCTAATT GCTCTTGTTA AACCTCGGAG CAAAGTGGTG TTATGC3GAG
 7801 ACCCAAGCA ATGCGGATTC TTCAATATGA TGCAGCTTAA GGTGAAGTTC AACCACAACA
 7861 TCTGCACTGA AGTATGTCTA AAAAGTATAT CCAGACGTTG CACGCGTCCA GTCACGCCCA
 7921 TCGTGTCTAC GTTGCACTAC GGAGGCAAGA TGCGCACGAC CAACCCGTGC AACAAACCCA
 7981 TAATCATAGA CACCACAGGA CAGACCAAGC CCAAGCCAGG AGACATCGTG TTAACATGCT
 8041 TCCGAGGCTG GGCAAAGCAG CTGCAGTTGG ACTACCGTGG ACACGAAGTC ATGACASCAG
 8101 CAGCATCTCA GGCCTCACCC GCGAAAGGGG TATACGCCGT AAGGCAGAAG GTGAAT3AAA
 8161 ATCCCTTGTA TGCCCTGCG TCGGAGCAGC TGAATGTACT GCTGACGCGC ACTGAG3ATA
 8221 GGCTGGTGTG GAAAACGCTG GCCGGCGATC CCTGGATTAA GGTCTATCA AACATTCCAC
 8281 AGGGTAACCT TACGGCCACA TTGGAAGAAT GGCAAGAAGA ACACGACAAA ATAATGAAGG
 8341 TGATTGAAGG ACCGGCTGCG CCTGTGGACG CGTTCAGAA CAAAGCGAAC GTGTGTGGG
 8401 CGAAAAGCCT GGTGCCTGTC CTGGACACTG CCGGAATCAG ATTGACAGCA GAGGAGTGGA
 8461 GCACCATAAT TACAGCATT AAGGAGGACA GAGCTTACTC TCCAGTGGTG GCCTTGAATG
 8521 AAATTTGCAC CAAGTACTAT GGAGTTGACC TGGACAGTGG CCTGTTTCT GCCCCGAAGG
 8581 TGTCCCTGTA TTACGAGAAC AACCCTGGG ATAACAGACC TGGTGAAGG ATGTAT3GAT
 8641 TCAATGCCGC AACAGCTGCC AGGCTGGAAG CTAGACATAC CTTCTGAAG GGGCAGTGGC
 8701 ATACGGGCAA GCAGGCAGTT ATCGCAGAAA GAAAAATCCA ACCGCTTCT GTGCTG3ACA
 8761 ATGTAATTCC TATCAACCGC AGGCTGCCGC ACGCCCTGGT GGCTGAGTAC AAGACG3TTA
 8821 AAGGCAGTAG GGTGAGTGG CTGGTCAATA AAGTAAGAGG GTACCAGTGC CTGCTG3TGA
 8881 GTGAGTACAA CCTGGCTTTG CCTCGACGCA GGGTCACTTG GTTGTACCG CTGAAT3TCA
 8941 CAGGCGCCGA TAGGTGCTAC GACCTAAGTT TAGGACTGCC GGCTGACGCC GGCAGT3TCG
 9001 ACTTGGTCTT TGTGAACATT CACACGGAAT TCAGAATCCA CCACTACCAG CAGTGT3TCG
 9061 ACCACGCCAT GAAGCTGCAG ATGCTTGGGG GAGATGCGCT ACGACTGCTA AAACCCGGCG
 9121 GCATCTTGAT GAGAGCTTAC GGATACGCCG ATAAAATCAG CGAAGCCGTT GTTCTCTCT
 9181 TAAGCAGAAA GTTCTCGTCT GCAAGAGTGT TGCGCCCGGA TTGTGTCAAC AGCAAT3ACAG
 9241 AAGTGTCTT GCTGTTCTCC AACTTTGACA ACGGAAAGAG ACCCTCTACG CTACAC3AGA
 9301 TGAATACCAA GCTGAGTGCC GTGTATGCCG GAGAAGCCAT GCACACGGCC GGGTGT3CAC
 9361 CATCCTACAG AGTTAAGAGA GCAGACATAG CCACGTGCAC AGAAGCGGCT GTGGTTA3ACG-

FIGURE 29A

67/240

9421 CAGCTAACGC CCGTGGAAC TTAGGGGATG GCGTATGCAG GGCCGTGGCG AAGAAATGGC
9481 CGTCAGCCTT TAAGGGAGCA GCAACACCAG TGGGCACAAT TAAACAGTC ATGTGCGGCT
9541 CGTACCCCGT CATCCACGCT GTAGCGCCTA ATTTCTCTGC CACGACTGAA GCGGAAGGGG
9601 ACCGCGAATT GGCCGCTGTC TACCGGGCAG TGGCCGCCGA AGTAAACAGA CTGTCACTGA
9661 GCAGCGTAGC CATCCCGCTG CTGTCCACAG GAGTGTTCAG CGGCGGAAGA GATAGGCTGC
9721 AGCAATCCCT CAACCATCTA TTCACAGCAA TGGACGCCAC GGACGCTGAC GTGACCATCT
9781 ACTGCAGAGA CAAAAGTTGG GAGAAGAAAA TCCAGGAAGC CATTGACATG AGGACGGCTG
9841 TGGAGTTGCT CAATGATGAC GTGGAGCTGA CCACAGACTT GGTGAGAGTG CACCCGGACA
9901 GCAGCCTGGT GGGTCGTAAG GGCTACAGTA CCACTGACGG GTCGCTGTAC TCGTACTTTG
9961 AAGGTACGAA ATTCAACCAG GCTGCTATTG ATATGGCAGA GATACTGACG TTGTGGCCCA
10021 GACTGCAAGA GGCAAACGAA CAGATATGCC TATACGCGCT GGGCGAAACA ATGGACAACA
10081 TCAGATCCAA ATGTCCCGTG AACGATTCCG ATTCATCAAC ACCTCCCAGG ACAGTGCCCT
10141 GCCTGTGCCG CTACGCAATG ACAGCAGAAC GGATCGCCCC CCTTAGGTCA CACCAAGTTA
10201 AAAGCATGGT GGTTCGCTCA TCTTTTCCCC TCCCGAAATA CCATGTAGAT GGGGTGCGA
10261 AGGTAAAGTG CGAGAAGGTT CTCCTGTTCTG ACCCGACGGT ACCTTCAGTG GTTAGTCCGC
10321 GGAAGTATGC CGCATCTACG ACGGACCACT CAGATCGGTC GTTACGAGGG TTTGACTTTG
10381 ACTGGACCAC CGACTCGTCT TCCACTGCCA GCGATACCAT GTCGCTACCC AGTTTGCACT
10441 CGTGTGACAT CGACTCGATC TACGAGCCAA TGGTCCCAT AGTAGTGACG GCTGACGTAC
10501 ACCTGAACC CGCAGGCATC GCGGACCTGG CCGCAGATGT GCACCCTGAA CCCGCAGACC
10561 ATGTGGACCT GGAGAACCCG ATTCCTCCAC CGCGCCCGAA GAGAGCTGCA TACCTTGCCCT
10621 CCCGCGCGGC GGAGCGACCG GTGCCGCGCG CGAGAAAGCC GACGCTGCC CCAAGGACTG
10681 CGTTTAGGAA CAAGCTGCCT TTGACGTTCTG CCGACTTTGA CGAGCACGAG TTCGATGCCG
10741 TGGCCTCCGG GATTACTTTC GGAGACTTCG ACGACGTCTT GCGACTAGGC CGCGCGGGTG
10801 CATATATTTT CTCCTCGGAC ACTGGCAGCG GACATTTACA ACAAATATCC GTTAGGCAGC
10861 ACAATCTCCA GTGCGCACAA CTGGATGCGG TCCAGGAGGA GAAATGTAC CCGCCAAAT
10921 TGATACTGA GAGGGAGAAG CTGTTGCTGC TGAATATGCA GATGCACCCA TCGGAGGCTA
10981 ATAAGAGTCG ATACCACTCT CGCAAAGTGG AGAACATGAA AGCCACGGTG GTGGACAGGC
11041 TCACATCGGC GGCCAGATTG TACACGGGAG CCGACGTAGG CCGCATACCA ACATACCGG
11101 TTCGGTACCC CCGCCCCGTG TACTCCCTA CCGTGATCGA AAGATTCTCA AGCCCCGATG
11161 TAGCAATCGC AGCGTGCAAC GAATACCTAT CCAGAAATTA CCAACAGTG GCGTCGTACC
11221 AGATAACAGA TGAATACGAC GCATACTTGG ACATGGTTGA CGGGTCGGAT AGTTGCTTTG
11281 ACAGAGCGAC ATTCTGCCCG GCGAAGCTCC GGTGCTACCC GAAACATCAT GCGTACCACC
11341 AGCCGACTGT ACGCAGTGCC GTCCCGTCAC CCTTTCAGAA CACACTACAG AACGTGCTAG
11401 CGGCTGCCAC CAAGAGAAAC TGCAACGTCA CGCAAATGCG AGAACTACCC ACCATGGACT
11461 CGGCAGTGTT CAACGTGGAG TGCTTCAAGC GCTATGCCTG CTCCGGAGAA TATTGGGAAG
11521 AATATGCTAA ACAACCTATC CGGATAACCA CTGAGAACAT CACTACCTAT GTGACCAAAT
11581 TGAAAGGCCC GAAAGCTGCT GCCTTGTTCTG CTAAGACCCA CAACTTGGTT CCGCTGCAGG
11641 AGGTTCCCAT GGACAGATTG ACGGTCGACA TGAACGAGA TGTCAAAGTC ACTCCAGGGA
11701 CGAAACACAC AGAGGAAAGA CCCAAGTCC AGGTAATTCA AGCAGCGGAG CCATTGGCGA
11761 CCGCTTACCT GTGCGGCATC CACAGGGAAT TAGTAAGGAG ACTAAATGCT GTGTTACGCC
11821 CTAACGTGCA CACATTGTTT GATATGTCGG CCGAAGACTT TGACGCGATC ATCGCCTCTC
11881 ACTTCCACCC AGGAGACCCG GTTCTAGAGA CGGACATTGC ATCATTGAC AAAAGCCAGG
11941 ACGACTCCTT GGCTCTTACA GGTTTAATGA TCCTCGAAGA TCTAGGGGTG GATCAGTACC
12001 TGCTGGACTT GATCGAGGCA GCCTTTGGGG AAATATCCAG CTGTCACTA CCAACTGGCA
12061 CGCGCTTCAA GTTCGGAGCT ATGATGAAAT CGGGCATGTT TCTGACTTTG TTTATTAACA
12121 CTGTTTTGAA CATCACCATA GCAAGCAGGG TACTGGAGCA GAGACTCACT GACTCCGCC
12181 GTGCGGCCTT CATCGGCGAC GACAACATCG TTCACGGAGT GATCTCCGAC AAGCTGATGG
12241 CGGAGAGGTG CGCGTCGTGG GTCAACATGG AGGTGAAGAT CATTGACGCT GTCATGGGGC
12301 AAAAACCCCC ATATTTTTGT GGGGGATTCA TAGTTTTTGA CAGCGTCACA CAGACCGCCT
12361 GCCGTGTTTC AGACCCACTT AAGCGCTGT TCAAGTTGGG TAAGCCGCTA ACAGCTGAAG
12421 ACAAGCAGGA CGAAGACAGG CGACGAGCAC TGAGTGACGA GGT

FIGURE 29E

68/240

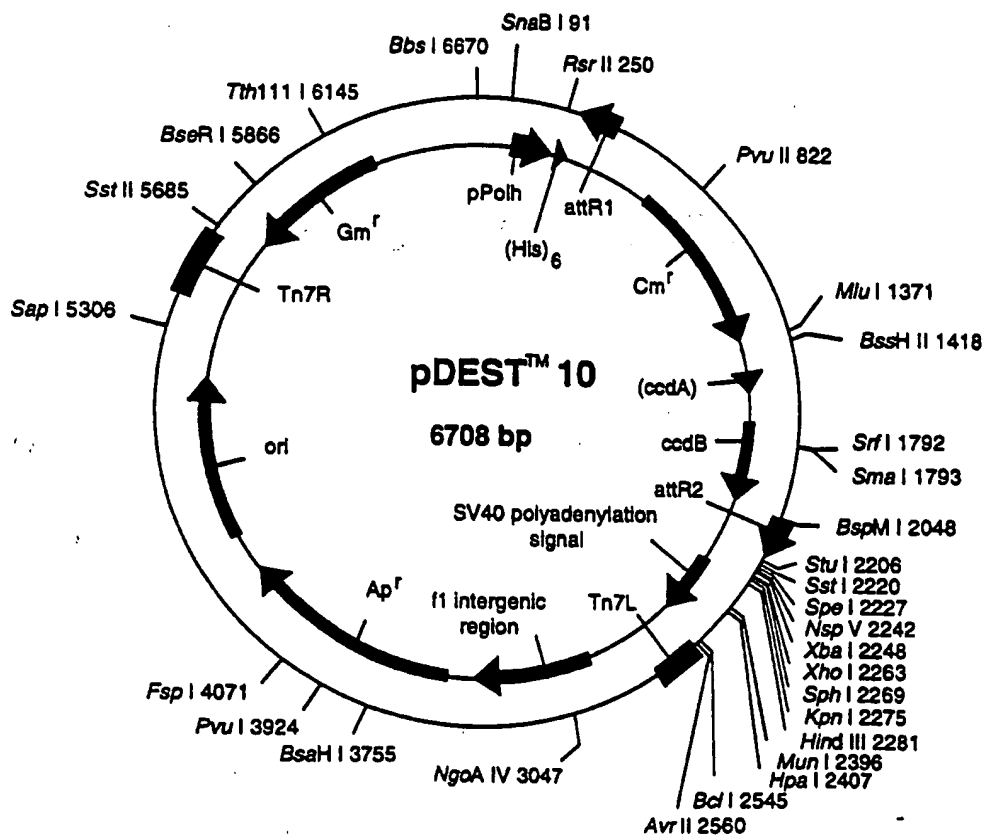
**Figure 30A: pDEST10 Polyhedron Promoter with N-His6,
Baculovirus Transfer Plasmid**

154 *← mRst from polyhedrin promoter*
 aaa taa gta ttt tac tgt ttt cgt aac agt ttt gta ata aaa aaa cct ata
 ttt att cat aaa atg aca aaa gca ttg tca aaa cat tat ttt ttt gga tat

205
 aat att ccg gat tat tca tac cgt ccc acc atc ggg cgc gga tct cgg tcc
 tta taa ggc cta ata agt atg gca ggg tgg tag ccc gcg cct aga gcc agg

256 *Met Ser Tyr Tyr His His His His His Asp Tyr Asp Ile Pro*
 gaa acc atg tgg tac tac cat cac cat cac cat cac gat tac gat atc cca
 ctt tgg tac agc atg atg gta gtg gta gtg gta gtg cta atg cta tag ggt

307 *TEV protease*
 Ttr Thr Glu Asn Leu Tyr Phe Gln Gly Ile Thr Ser Leu Tyr Lys Lys
 acg acc gaa aac ctg tat ttt cag ggc atc aca agt tgg tgc ada aaa ggc
 tgc tgg ctt ttg gac ata aaa gtc ccg tag tgt tca aac atg ttt ttc oga
 attR1 Int



69/240

pDEST10 6708 bp

<u>Location (Base Nos.)</u>	<u>Gene Encoded</u>
23..152	Ppolh
461..337	attR1
711..1370	CmR
1490..1574	inactivated ccdA
1712..2017	ccdB
2058..2182	attR2
3394..4369	ampR
4510..5164	ori
5658..62	genR

```

1  CCCC GGATGA AGTGGTTCGC ATCCTCGGTT TTCTGGAAGG CGAGCATCGT TTGTTGCCCC
61 AGGACTCTAG CTATAGTTCT AGTGGTTGGC TACGTATACT CCGGAATATT AATAGATCAT
121 GGAGATAATT AAAATGATAA CCATCTCGCA AATAAATAAG TATTTTACTG TTTTCGTAAC
181 AGTTTTGTAA TAAAAAACC TATAAATATT CCGGATTATT CATACCGTCC CACCATCGGG
241 CGCGGATCTC GGTCCGAAAC CATGTCGTAC TACCATCACC ATCACCATCA CGATTACGAT
301 ATCCCAACGA CCGAAACCT GTATTTTCAG GGCATCACAA GTTTGTACAA AAAAGCTGAA
361 CGAGAAACGT AAAATGATAT AAATATCAAT ATATTAAATT AGATTTTGCA TAAAAACAG
421 ACTACATAAT ACTGTAAAAC ACAACATATC CAGTCACTAT GGCGGCCGCT AAGTTGGCAG
481 CATCACCCGA CGCACTTTGC GCCGAATAAA TACCTGTGAC GGAAGATCAC TTCGCAGAAT
541 AAATAAATCC TGGTGTCCCT GTTGATACCG GGAAGCCCTG GGCCAACTTT TGGCGAAAAT
601 GAGACGTGTA TCGGCACGTA AGAGGTTCCA ACTTTCACCA TAATGAAATA AGATCACTAC
661 CGGGCGTATT TTTTGAGTTA TCGAGATTTT CAGGAGCTAA GGAAGCTAAA ATGGAGAAAA
721 AAATCACTGG ATATACCACC GTTGATATAT CCCAATGGCA TCGTAAAGAA CATTTTGAGG
781 CATTTTCAGT AGTTGCTCAA TGTACCTATA ACCAGACCGT TCAGCTGGAT ATTACGGCCT
841 TTTTAAAGAC CGTAAAGAAA AATAAGCACA AGTTTTATCC GGCTTTTATT CACATTCTTG
901 CCCGCCGTGAT GAATGCTCAT CCGGAATTCC GTATGGCAAT GAAAGACGGT GAGCTGGTGA
961 TATGGGATAG TGTTCAACCCT TGTTACACCG TTTTCCATGA GCAAACCTGAA ACGTTTTCAT
1021 CGCTCTGGAG TGAATACCAC GACGATTTCG GGCAGTTTCT ACACATATAT TCGCAAGATG
1081 TGGCGTGTGA CCGTGAAAAC CTGGCCTATT TCCCTAAAGG GTTTATTGAG AATATGTTTT
1141 TCGTCTCAGC CAATCCCTGG GTGAGTTTCA CCAGTTTGA TTTAAACGTG GCCAATATGG
1201 ACAACTTCTT CGCCCCGTT TTCACCATGG GCAAATATTA TACGCAAGGC GACAAGGTGC
1261 TGATGCCGCT GGCGATTTCAG GTTCATCATG CCGTCTGTGA TGGCTTCCAT GTCGGCAGAA
1321 TGCTTAATGA ATTACAACAG TACTGCGATG AGTGGCAGGG CGGGGCGTAA ACCCGTGGAT
1381 CCGGCTTACT AAAAGCCAGA TAACAGTATG CGTATTTGCG CGCTGATTTT TCGCGTATAA
1441 GAATATATAC TGATATGTAT ACCCGAAGTA TGTCAAAAAG AGGTGTGCTA TGAAGCAGCG
1501 TATTACAGTG ACAGTTGACA GCGACAGCTA TCAGTTGCTC AAGGCATATA TGATGTCAAT
1561 ATCTCCGGTC TGGTAAGCAC AACCATGCAG AATGAAGCCC GTCGTCTGCG TGCCGAACGC
1621 TGGAAAGCGG AAAATCAGGA AGGGATGGCT GAGGTCGCCC GGTTTATTGA AATGAACGGC
1681 TCTTTTGTG ACGAGAACAG GGACTGGTGA AATGCAGTTT AAGGTTTACA CCTATAAAG
1741 AGAGAGCCGT TATCGTCTGT TTGTGGATGT ACAGAGTGAT ATTATTGACA CGCCCGGGCG
1801 ACGGATGGTG ATCCCCCTGG CCAGTGCACG TCTGCTGTCA GATAAAGTCT CCCGTGAAC
1861 TTACCCGGTG GTGCATATCG GGGATGAAAG CTGGCGCATG ATGACCACCG ATATGGCCAG
1921 TGTGCCGGTC TCCGTTATCG GGAAGAAGT GGCTGATCTC AGCCACCGCG AAAATGACAT
1981 CAAAAACGCC ATTAACCTGA TGTCTGGGG AATATAAATG TCAGGCTCCC TTATACACAG
2041 CCAGTCTGCA GGTCGACCAT AGTGAAGTGA TATGTTGTGT TTTACAGTAT TATGTAGTCT
2101 GTTTTTTATG CAAAATCTAA TTAAATATAT TGATATTTAT ATCATTTTAC GTTTCTCGTT
2161 CAGCTTTCTT GTACAAAGTG GTGATGCCAT GGATCCGGAA TTCAAAGGCC TACGTCGACG
2221 AGCTCAACTA GTGCGGCCGC TTTCGAATCT AGAGCCTGCA GTCTCGAGGC ATGCGGTACC
2281 AAGCTTGTG AGAAGTACTA GAGGATCATA ATCAGCCATA CCACATTTGT AGAGGTTTTA
2341 CTTGCTTTAA AAAACCTCCC ACACCTCCCC CTGAACCTGA AACATAAAAT GAATGCAATT
2401 GTTGTTGTGA ACTTGTTTAT TGCAGCTTAT AATGGTTACA AATAAAGCAA TAGCATCACA
2461 AATTTACAA ATAAAGCATT TTTTCACTG CATTCTAGTT GTGTTTGTG CAAACTCATC
2521 AATGTATCTT ATCATGTCTG GATCTGATCA CTGCTTGAGC CTAGGAGATC CGAACCAGAT
2581 AAGTGAAATC TAGTTCCAAA CTATTTTGTC ATTTTAAATT TTCGTATTAG CTTACGACGC-

```

FIGURE 30B

70/240

2641 TACACCCAGT TCCCATCTAT TTTGTCACTC TTCCCTAAAT AATCCTTAAA AACTCCATTT
2701 CCACCCCTCC CAGTTCCTCA CTATTTTGTC CGCCACAGC GGGGCATTTT TCTTCCTGTT
2761 ATGTTTTTAA TCAAACATCC TGCCAACCTC ATGTGACAAA CCGTCATCTT CGGCTACTTT
2821 TTCTCTGTCA CAGAATGAAA ATTTTCTGT CATCTCTTCG TTATTAATGT TTGTAATTGA
2881 CTGAATATCA ACGCTTATTT GCAGCCTGAA TGGCGAATGG GACGCGCCCT GTAGCGGCGC
2941 ATTAAGCGCG GCGGGTGTGG TGGTTACGCG CAGCGTGACC GCTACACTTG CCAGCGCCCT
3001 AGCGCCCGCT CCTTTCGCTT TCTTCCCTTC CTTTCTCGCC ACGTTCGCGG GCTTTCCTCCG
3061 TCAAGCTCTA AATCGGGGGC TCCCTTTAGG GTTCCGATTT AGTGCTTTAC GGCACCTCGA
3121 CCCCCAAAAA CTTGATTAGG GTGATGGTTC ACGTAGTGGG CCATCGCCCT GATAGACGGT
3181 TTTTCGCCCT TTGACGTTGG AGTCCACGTT CTTTAATAGT GGACTCTTGT TCCAAACTGG
3241 AACAACACTC AACCCTATCT CGGTCTATTC TTTTGATTTA TAAGGGATTT TGCCGATTTT
3301 GGCCTATTGG TTAAAAATG AGCTGATTTA AAAAAAATT AACCGGAATT TTAACAAAAT
3361 ATTAACGTTT ACAATTTTCA GTGGCATTTC TCGGGGAAAT GTGCGCGGAA CCCCTATTTG
3421 TTTATTTTTT TAAATACATT CAAATATGTA TCCGCTCATG AGACAATAAC CCTGATAAAT
3481 GCTTCAATAA TATTGAAAAA GGAAGAGTAT GAGTATTCAA CATTTCCGTG TCGCCCTTAT
3541 TCCCTTTTTT GCGGCATTTT GCCTTCCTGT TTTTGCTCAC CCAGAAACGC TGGTGAAAGT
3601 AAAAGATGCT GAAGATCAGT TGGGTGCACG AGTGGGTAC ATCGAACTGG ATCTCAACAG
3661 CGGTAAGATC CTTGAGAGTT TTCGCCCCGA AGAACGTTTT CCAATGATGA GCATTTTTAA
3721 AGTTCTGCTA TGTGGCGCGG TATTATCCCG TATTGACGCC GGGCAAGAGC AACTCGGTGCG
3781 CCGCATACAC TATTCTCAGA ATGACTTGGT TGAGTACTCA CCAGTCACAG AAAAGCATCT
3841 TACGGATGGC ATGACAGTAA GAGAATTATG CAGTGCTGCC ATAACCATGA GTGATAACAC
3901 TGCGGCCAAC TTACTTCTGA CAACGATCGG AGGACCGAAG GAGCTAACCG CTTTTTTGCA
3961 CAACATGGGG GATCATGTAA CTCGCCTTGA TCGTTGGGAA CCGGAGCTGA ATGAAGCCAT
4021 ACCAAACGAC GAGCGTGACA CCACGATGCC TGTAGCAATG GCAACAACGT TGCGCAAAC
4081 ATTAAGTGGC GAACTACTTA CTCTAGCTTC CCGGCAACAA TTAATAGACT GGATGGAGGC
4141 GGATAAAGTT GCAGGACCAC TTCTGCGCTC GGCCCTTCCG GCTGGCTGGT TTATTGCTGA
4201 TAAATCTGGA GCCGGTGAGC GTGGGTCTCG CGGTATCATT GCAGCACTGG GGCCAGATGG
4261 TAAGCCCTCC CGTATCGTAG TTATCTACAC GACGGGGAGT CAGGCAACTA TGGATGAACG
4321 AAATAGACAG ATCGCTGAGA TAGGTGCCTC ACTGATTAAG CATTTGGTAA TGTCAGACCA
4381 AGTTTACTCA TATATACTTT AGATTGATTT AAAAATTCAT TTTTAATTTA AAAGGATCTA
4441 GGTGAAGATC CTTTTTGATA ATCTCATGAC CAAATCCCT TAACGTGAGT TTTTCGTTCCA
4501 CTGAGCGTCA GACCCCGTAG AAAAGATCAA AGGATCTTCT TGAGATCCTT TTTTTCGCG
4561 CGTAATCTGC TGCTTGCAAA CAAAAAACC ACCGCTACCA GCGGTGGTTT GTTTGCCGGA
4621 TCAAGAGCTA CCAACTCTTT TTCCGAAGGT AACTGGCTTC AGCAGAGCGC AGATACCAAA
4681 TACTGTCTTT CTAGTGTAGC CGTAGTTAGG CCACCACTTC AAGAACTCTG TAGCACCGCC
4741 TACATACCTC GCTCTGCTAA TCCTGTACC AGTGGCTGCT GCCAGTGGCG ATAAGTCGTG
4801 TCTTACCGGG TTGGACTCAA GACGATAGTT ACCGGATAAG GCGCAGCGGT CGGGCTGAAC
4861 GGGGGGTTCG TGCACACAGC CCAGCTTGGA GCGAACGACC TACACCGAAC TGAGATACCT
4921 ACAGCGTGTAG CATTGAGAAA GCGCCACGCT TCCCGAAGGG AGAAAGGCGG ACAGGTATCC
4981 GGTAAGCGGC AGGGTCGGAA CAGGAGGAGG CACGAGGGG CTTCCAGGGG GAAACGCTG
5041 GTATCTTTAT AGTCTGTGCG GGTTCGCCA CCTCTGACTT GAGCGTCGAT TTTTGTGATG
5101 CTCGTACAGG GGGCGGAGCC TATGGAAAAA CGCCAGCAAC GCGGCCTTTT TACGGTTCCT
5161 GGCCTTTTGC TGGCCTTTTG CTCACATGTT CTTTCTGCG TTATCCCCTG ATTCTGTGGA
5221 TAACCGTATT ACCGCTTTG AGTGAGCTGA TACCGCTCGC CGCAGCCGAA CGACCGAGCG
5281 CAGCGAGTCA GTGAGCGAGG AAGCGGAAGA GCGCTGATG CCGTATTTTC TCCTTACGCA
5341 TCTGTGCGGT ATTTACACAC GCAGACCAGC CGCGTAACCT GGCAAAATCG GTTACGGTTG
5401 AGTAATAAAT GGATGCCCTG CGTAAGCGGG TGTGGGCGGA CAATAAAGTC TTAACTGAA
5461 CAAAATAGAT CTAAACTATG ACAATAAAGT CTTAAACTAG ACAGAATAGT TGTAAACTGA
5521 AATCAGTCCA GTTATGCTGT GAAAAAGCAT ACTGGACTTT TGTTATGGCT AAAGCAAAC
5581 CTTCAATTTT TGAAGTGCAA ATTGCCCCTG GTATTAAAGA GGGGCGTGGC CAAGGGCATG
5641 GTAAAGACTA TATTCGCGGC GTTGTGACAA TTTACCGAAC AACTCCGCGG CCGGGAAGCC
5701 GATCTCGGCT TGAACGAATT GTTAGGTGGC GGTACTTGGG TCGATATCAA AGTGATCAC
5761 TTCTTCCCGT ATGCCCACT TTGTATAGAG AGCCACTGCG GGATCGTCAC CGTAATCTGC
5821 TTGCACGTAG ATCACAATAG CACCAAGCGC GTTGGCCTCA TGCTTGAGGA GATTGATGAG
5881 CGCGGTGGCA ATGCCCTGCC TCCGGTGCTC GCGGAGACT GCGAGATCAT AGATATAGAT
5941 CTCACTACGC GGCTGCTCAA ACCTGGGCGA AACGTAAGCC GCGAGAGCGC CAACAACCGC
6001 TTCTTGGTGC AAGGCAGCAA GCGCGATGAA TGTCTTACTA CGGAGCAAGT TCCCGAGGTA
6061 ATCGGAGTCC GGCTGATGTT GGGAGTAGGT GGCTACGTCT CCGAACTCAC GACCGAAAAG-

FIGURE 30C

6121 ATCAAGAGCA GCCCGCATGG ATTTGACTTG GTCAGGGCCG AGCCTACATG TGCGAATGAT
6181 GCCCATACTT GAGCCACCTA ACTTTGTTTT AGGGCGACTG CCCTGCTGCG TAACATCGTT
6241 GCTGCTGCGT AACATCGTTG CTGCTCCATA ACATCAAACA TCGACCCACG GCGTAACGCG
6301 CTTGCTGCTT GGATGCCCCG GGCATAGACT GTACAAAAAA ACAGTCATAA CAAGCCATGA
6361 AAACCGCCAC TGCGCCGTTA CCACCGCTGC GTTCGGTCAA GGTTCCTGGAC CAGTTGCGTG
6421 AGCGCATACG CTACTTGCAT TACAGTTTAC GAACCGAACA GGCTTATGTC AACTGGGTTT
6481 GTGCCTTCAT CCGTTTCCAC GGTGTGCGTC ACCCGGCAAC CTTGGGCAGC AGCGAAGTCG
6541 AGGCATTCTT GTCCTGGCTG GCGAACGAGC GCAAGGTTTC GGTCTCCACG CATCGTCAGG
6601 CATTGGCGGC CTTGCTGTTC TTCTACGGCA AGGTGCTGTG CACGGATCTG CCCTGGCTTC
6661 AGGAGATCGG AAGACCTCGG CCGTCGCGGC GCTTGCCGGT GGTGCTGA

72/240

Figure 31A:

pDEST11

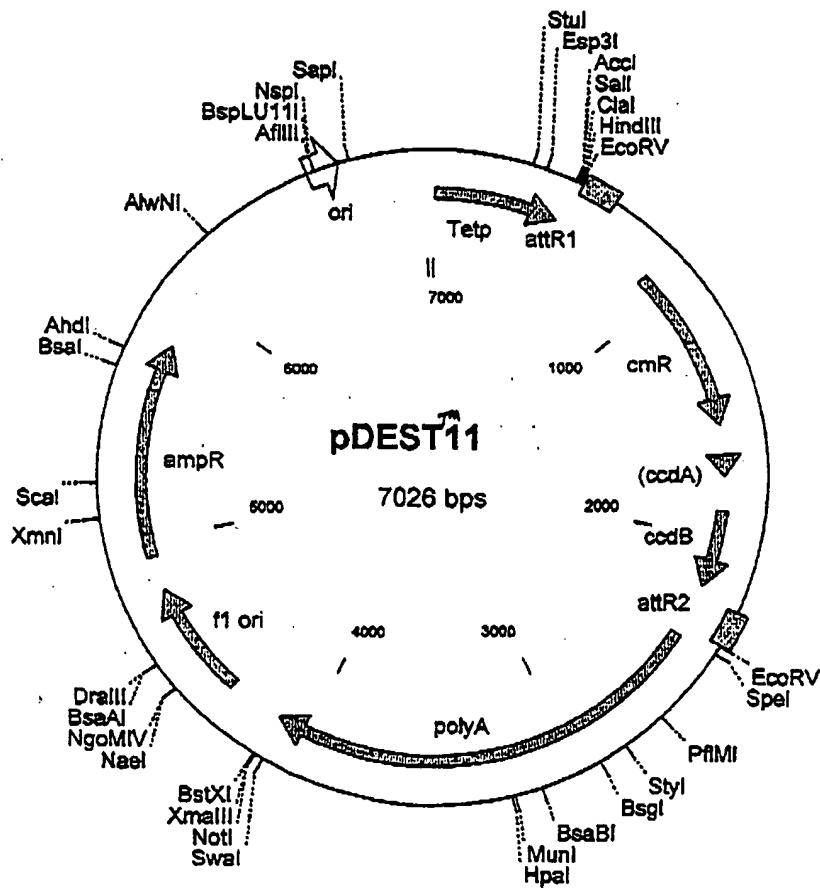
Tet-regulated eukaryotic expression

358 tag tga acc ~~gfc~~ ^{mRNA from CMV promoter (controlled by tetracycline)} aga tgg cct gga gac gcc atc cac gct gtt tgg acc tcc
 atc act tgg cag tct agc gga cct ctg cgg tag gtg cga caa aac tgg agg

409 ata gaa gac acc ggg acc gat cca gcc tcc gcg gcc cgg aat tgg agc tgg
 tat ctt ctg tgg ccc tgg cta ggt cgg agg cgc cgg ggc tta agc tgg agc

460 gta ccc ggg gat cct cta gag tgg ^{Sal}agg ^{Cla}tgg ^{Hind3}acg gta ^{EcoRV}tgg ^{EcoRV}ata ^{EcoRV}agg ^{EcoRV}tgg ^{EcoRV}ata
 cat ggg ccc cta gga gat ctc agc tcc agc tgc cat agc tat tgg agc tat

511 tca ~~aca agt tgg~~ ^{Int} ~~aga aac aac~~ ^{attR1} ~~gct gaa cga gaa acg taa~~ ~~aat gat ata gat~~
 agt ~~tgt tca aac atg ttt~~ ~~tct cga ctt gct ctc tgc att tta cta cat tta~~



73/240

pDEST11 7026 bp

<u>Location (Base Nos.)</u>		<u>Gene Encoded</u>
4...479		Tetp ((Tet operator) 7 and min hCMV promoter)
638...514		attR1
888...1547		CmR
1667...1751		inactivated ccdA
1889...2194		ccdB
2235...2359		attR2
2402...4132		polyA
4347...4803		f1 ori
4940...5797		ampR

1	CGAGTTTACC	ACTCCCTATC	AGTGATAGAG	AAAAGTGAAA	GTCGAGTTTA	CCACTCCCTA
61	TCAGTGATAG	AGAAAAGTGA	AAGTCGAGTT	TACCACTCCC	TATCAGTGAT	AGAGAAAAGT
121	GAAAGTCGAG	TTTACCACTC	CCTATCAGTG	ATAGAGAAAA	GTGAAAGTCG	AGTTTACCAC
181	TCCCTATCAG	TGATAGAGAA	AAGTGAAAGT	CGAGTTTACC	ACTCCCTATC	AGTGATAGAG
241	AAAAGTGAAA	GTCGAGTTTA	CCACTCCCTA	TCAGTGATAG	AGAAAAGTGA	AAGTCGAGCT
301	CGGTACCCGG	GTCGAGTAGG	CGTGTACGGT	GGGAGGCCTA	TATAAGCAGA	GCTCGTTTAG
361	TGAACCGTCA	GATCGCCTGG	AGACGCCATC	CACGCTGTTT	TGACCTCCAT	AGAAGACACC
421	GGGACCGATC	CAGCCTCCGC	GGCCCCGAAT	TCGAGCTCGG	TACCCGGGGA	TCCTCTAGAG
481	TCGAGGTCGA	CGGTATCGAT	AAGCTTGATA	TCAACAAGTT	TGTACAAAAA	AGCTGAACGA
541	GAAACGTAAA	ATGATATAAA	TATCAATATA	TTAAATTAGA	TTTTGCATAA	AAAACAGACT
601	ACATAATACT	GTAATAACACA	ACATATCCAG	TCACTATGGC	GGCCGCTAAG	TTGGCAGCAT
661	CACCCGACGC	ACTTTGCGCC	GAATAAATAC	CTGTGACGGA	AGATCACTTC	GCAGAATAAA
721	TAAATCCTGG	TGTCCCTGTT	GATACCGGGA	AGCCCTGGGC	CAACTTTTGG	CGAAAATGAG
781	ACGTTGATCG	GCACGTAAGA	GGTTCCAAC	TTCAACATAA	TGAAATAAGA	TCACTACCGG
841	CGGTATTTTT	TGAGTTATCG	AGATTTTCAG	GAGCTAAGGA	AGCTAAAATG	GAGAAAAAAA
901	TCCTGGGATA	TACCACCGTT	GATATATCCC	AATGGCATCG	TAAAGAACAT	TTTGAGGCAT
961	TTCACTCAGT	TGCTCAATGT	ACCTATAACC	AGACCGTTCA	GCTGGATATT	ACGGCCTTTT
1021	TAAAGACCGT	AAAGAAAAAT	AAGCACAAGT	TTTATCCGGC	CTTTATTAC	ATTCTTGCCC
1081	GCCTGATGAA	TGCTCATCCG	GAATCCGTA	TGGCAATGAA	AGACGGTGAG	CTGGTGATAT
1141	GGGATAGTGT	TCACCCTTGT	TACACCGTTT	TCCATGAGCA	AACTGAAACG	TTTTTCATCGC
1201	TCTGGAGTGA	ATACCACGAC	GATTTCCGGC	AGTTTCTACA	CATATATTCG	CAAGATGTGG
1261	CGTGTTACGG	TGAAAACCTG	GCCTATTTCC	CTAAAGGGTT	TATTGAGAAT	ATGTTTTTCG
1321	TCTCAGCCAA	TCCCTGGGTG	AGTTTCACCA	GTTTTGATTT	AAACGTGGCC	AATATGGACA
1381	ACTTCTTCGC	CCCCGTTTTT	ACCATGGGCA	AATATTATAC	GCAAGGCGAC	AAGGTGCTGA
1441	TGCCGCTGGC	GATTCAGGTT	CATCATGCCG	TCTGTGATGG	CTTCCATGTC	GGCAGAATGC
1501	TTAATGAATT	ACAACAGTAC	TGCGATGAGT	GGCAGGGCGG	GGCGTAAAGA	TCTGGATCCG
1561	GCTTACTAAA	AGCCAGATAA	CAGTATGCGT	ATTTGCGCGC	TGATTTTTTC	GGTATAAGAA
1621	TATATACTGA	TATGTATACC	CGAAGTATGT	CAAAAAGAGG	TGTGCTATGA	AGCAGCGTAT
1681	TACAGTGACA	GTTGACAGCG	ACAGCTATCA	GTTGCTCAAG	GCATATATGA	TGTCAATATC
1741	TCCGGTCTGG	TAAGCACAA	CATGCAGAAT	GAAGCCCGTC	GTCTGCGTGC	CGAACGCTGG
1801	AAAGCGGAAA	ATCAGGAAGG	GATGGCTGAG	GTGCCCCGGT	TTATTGAAAT	GAACGGCTCT
1861	TTTGCTGACG	AGAACAGGGA	CTGGTGAAAT	GCAGTTTAAG	GTTTACACCT	ATAAAAAGAGA
1921	GAGCCGTTAT	CGTCTGTTTG	TGGATGTACA	GAGTGATATT	ATTGACACGC	CCGGGCGACG
1981	GATGGTGATC	CCCCTGGCCA	GTGCACGTCT	GCTGTGAGAT	AAAGTCTCCC	GTGAACCTTA
2041	CCCGGTGGTG	CATATCGGGG	ATGAAAGCTG	GCGCATGATG	ACCACCGATA	TGGCCAGTGT
2101	GCCGGTCTCC	GTTATCGGGG	AAGAAGTGGC	TGATCTCAGC	CACCGCGAAA	ATGACATCAA
2161	AAACGCCATT	AACCTGATGT	TCTGGGGAAT	ATAAATGTCA	GGCTCCCTTA	TACACAGCCA
2221	GTCTGCAGGT	CGACCATAGT	GACTGGATAT	GTTGTGTTTT	ACAGTATTAT	GTAGTCTGTT
2281	TTTTATGCAA	AATCTAATTT	AATATATTGA	TATTTATATC	ATTTTACGTT	TCTCGTTTAC
2341	CTTTCTTGTA	CAAAGTGGTT	GATATCGAAT	TCCTGCAGCC	CGGGGGATCC	ACTAGTTCTA
2401	GAGCACTGCG	ATGAGTGGCA	GGGCGGGGCG	TAATTTTTTT	AAGGCAGTTA	TTGGTGGCCT
2461	TAAACGCCTG	GTGCTACGCC	TGAATAAGTG	ATAATAAGCG	GATGAATGGC	AGAAATTTCG
2521	CGGATCTTTG	TGAAGGAACC	TTACTTCTGT	GGTGTGACAT	AATTGGACAA	ACTACCTACA-

FIGURE 31B

2581 GAGATTTTAA GCTCTAAGGT AAATATAAAA TTTTAAAGTG TATAATGTGT TAAACTACTG
2641 ATTCTAATTG TTTGTGTATT TTAGATTCCA ACCTATGGAA CTGATGAATG GGAGCAGTGG
2701 TGGAAATGCCT TTAATGAGGA AAACCTGTTT TGCTCAGAAG AAATGCCATC TAGTGATGAT
2761 GAGGCTACTG CTGACTCTCA ACATTCTACT CCTCCAAAAA AGAAGAGAAA GGTAGAAGAC
2821 CCCAAGGACT TTCCTTCAGA ATTGCTAAGT TTTTGTAGTC ATGCTGTGTT TAGTAATAGA
2881 ACTCTTGCTT GCTTTGCTAT TTACACCACA AAGGAAAAAG CTGCACTGCT ATACAAGAAA
2941 ATTATGGAAA AATATTCTGT AACCTTTATA AGTAGGCATA ACAGTTATAA TCATAACATA
3001 CTGTTTTTTC TTAATCCACA CAGGCATAGA GTGTCTGCTA TTAATAACTA TGCTCAAAAA
3061 TTGTGTACCT TTAGCTTTT AATTTGTAAA GGGGTTAATA AGGAATATTT GATGTATAGT
3121 GCCTTGACTA GAGATCATAA TCAGCCATAC CACATTTGTA GAGGTTTTAC TTGCTTTAAA
3181 AAACCTCCCA CACCTCCCC TGAACCTGAA ACATAAAATG AATGCAATTG TTGTTGTTAA
3241 CTTGTTTATT GCAGCTTATA ATGGTTACAA ATAAAGCAAT AGCATCACAA ATTTACAAAA
3301 TAAAGCATTT TTTTCACTGC ATTCTAGTTG TGGTTGTCC AAACCTATCA ATGTATCTTA
3361 TCATGTCTGG ATCCCCAGGA AGCTCCTCTG TGCTCTCATA AACCTTAACC TCCTCTACTT
3421 GAGAGGACAT TCCAATCATA GGCTGCCCAT CCACCCTCTG TGCTCTCTG TTAATTAGGT
3481 CACTTAACAA AAAGGAAATT GGGTAGGGT TTTTACAGA CCGCTTTCTA AGGGTAATTT
3541 TAAAAATATCT GGGAAAGTCCC TTCCACTGCT GTGTTCCAGA AGTGTGTTGTA AACAGCCCAC
3601 AAATGTCAAC AGCAGAAACA TACAAGCTGT CAGCTTTGCA CAAGGGCCCA ACACCCTGCT
3661 CATCAAGAAG CACTGTGGTT GCTGTGTTAG TAATGTGCAA AACAGGAGGC ACATTTTCCC
3721 CACCTGTGTA GGTTCACAAA TATCTAGTGT TTTTCAATTTT ACTTGGATCA GGAACCCAGC
3781 ACTCCACTGG ATAAGCATT TCCCTATCCA AAACAGCCTT GTGGTCAGTG TTCATCTGCT
3841 GACTGTCAAC TGTAGCATTT TTTGGGGTTA CAGTTTGAGC AGGATATTTG GTCCTGTAGT
3901 TTGCTAACAC ACCCTGCAGC TCCAAAGGTT CCCCACCAAC AGCAAAAAAA TGAAAAATTG
3961 ACCCTTGAAT GGGTTTTCCA GCACCATTTT CATGAGTTTT TTGTGTCCCT GAATGCAAGT
4021 TTAACATAGC AGTTACCCCA ATAACCTCAG TTTTAAACAGT AACAGCTTCC CACATCAAAA
4081 TATTTCCACA GGTAAAGTCC TCATTTAAAT TAGGCAAAGG AATTGCTCTA GAGCGGCCGC
4141 CACCGCGGTG GAGCTCCAAT TCGCCCTATA GTGAGTCGTA TTACGCGCGC TCACTGGCCG
4201 TCGTTTTACA ACGTCGTGAC TGGGAAAACC CTGGCGTTAC CCAACTTAAT CGCCTTGACG
4261 CACATCCCCC TTTCGCCAGC TGGCGTAATA GCGAAGAGGC CCGCACCAGT CGCCCTTCCC
4321 AACAGTTGCG CAGCCTGAAT GCGGAATGGG ACGCGCCCTG TAGCGGCGCA TTAAGCGCGG
4381 CGGTTGTGGT GGTACGCGC AGCGTGACCG CTACACTTGC CAGCGCCCTA GCGCCGCTC
4441 CTTTCGCTTT CTTCCCTTCC TTTCTCGCCA CGTTCCGCGG CTTTCCCGCT CAAGCTCTAA
4501 ATCGGGGGCT CCCTTTAGGG TTCCGATTTA GTGCTTTACG GCACCTCGAC CCCAAAAAAC
4561 TTGATTAGGG TGATGGTTCA CGTAGTGGGC CATCGCCCTG ATAGACGGTT TTTGCGCCCTT
4621 TGACGTTGGA GTCCACGTTT TTTAATAGTG GACTCTTGTT CCAAACCTGGA ACAACACTCA
4681 ACCCTATCTC GGTCTATTCT TTTGATTTAT AAGGGATTTT GCCGATTTCT GCCTATTGGT
4741 TAAAAAATGA GCTGATTTAA CAAAAATTTA ACGCGAATTT TAACAAAATA TTAACGCTTA
4801 CAATTTAGGT GGCACTTTTT GGGGAAATGT GCGCGGAACC CCTATTTGTT TATTTTCTA
4861 AATAGTATCA AATATGTATC CGCTCATGAG ACAATAACCC TGATAAATGC TTCAATAATA
4921 TTGAAAAAGG AAGAGTATGA GTATTCAACA TTTCCGTGTC GCCCTTATTC CCTTTTTTGC
4981 GGCATTTTGC CTTCTGTGTT TTGCTCAGCC AGAAACGCTG GTGAAAGTAA AAGATGCTGA
5041 AGATCAGTTG GGTGCACGAG TGGGTTACAT CGAACTGGAT CTCAACAGCG GTAAGATCCT
5101 TGAGAGTTTT CGCCCCGAAG AACGTTTTTC AATGATGAGC ACTTTTAAAG TTCTGCTATG
5161 TGGCGCGGTA TTATCCCGTA TTGACGCCGG GCAAGAGCAA CTCGGTCGCC GCATACACTA
5221 TTCTCAGAA GACTTGTTG AGTACTCACC AGTCACAGAA AAGCATCTTA CGGATGGCAT
5281 GACAGTAAGA GAATTATGCA GTGCTGCCAT AACCATGAGT GATAACACTG CGGCCAACTT
5341 ACTTCTGACA ACGATCGGAG GACCGAAGGA GCTAACCGCT TTTTGCACA ACATGGGGGA
5401 TCATGTAAC CGCCTTGATC GTTGGGAACC GGAGCTGAAT GAAGCCATAC CAAACGACGA
5461 GCGTGACACC ACGATGCCTG TAGCAATGGC AACAACGTTG CGCAAACCTAT TAACTGGCGA
5521 ACTACTTACT CTAGCTTCCC GGCAACAATT AATAGACTGG ATGGAGGCGG ATAAAGTTGC
5581 AGGACCACCT CTGCGCTCGG CCCTTCCGGC TGGCTGGTTT ATTGCTGATA AATCTGGAGC
5641 CGGTGAGCGT GGTCTCGCG GTATCATTTG AGCACTGGGG CCAGATGGTA AGCCCTCCCG
5701 TATCGTAGTT ATCTACACGA CGGGAGTCA GGCAACTATG GATGAACGAA ATAGACAGAT
5761 CGCTGAGATA GGTGCCTCAC TGATTAGCA TTGGTAACTG TCAGACCAAG TTTACTCATA
5821 TATACTTTAG ATTGATTTAA AACTTCATTT TTAATTTAAA AGGATCTAGG TGAAGATCCT
5881 TTTTGATAAT CTCATGACCA AAATCCCTTA ACGTGAGTTT TCGTTCCACT GAGCGTCAGA
5941 CCCCCTAGAA AAGATCAAAG GATCTTCTTG AGATCCTTTT TTTCTGCGCG TAATCTGCTG
6001 CTTGCAACAA AAAAAACCAC CGCTACCAGC GGTGGTTTGT TTGCCGGATC AAGAGCTACC-

75/240

6061 AACTCTTTTT CCGAAGGTAA CTGGCTTCAG CAGAGCGCAG ATACCAAATA CTGTCCTTCT
6121 AGTGTAGCCG TAGTTAGGCC ACCACTTCAA GAACTCTGTA GCACCGCCTA CATACCTCGC
6181 TCTGCTAATC CTGTTACCAG TGGCTGCTGC CAGTGGCGAT AAGTCGTGTC TTACCGGGTT
6241 GGA CTCAAGA CGATAGTTAC CGGATAAGGC GCAGCGGTCG GGCTGAACGG GGGGTTCTGT
6301 CACACAGCCC AGCTTGGAGC GAACGACCTA CACCGAACTG AGATACCTAC AGCGTGAGCT
6361 ATGAGAAAGC GCCACGCTTC CCGAAGGGAG AAAGGCGGAC AGGTATCCGG TAAGCGGCAG
6421 GGTCGGAACA GGAGAGCGCA CGAGGGAGCT TCCAGGGGGA AACGCCTGGT ATCTTTATAG
6481 TCCTGTCGGG TTTCGCCACC TCTGACTTGA GCGTCGATTT TTGTGATGCT CGTCAGGGGG
6541 GCGGAGCCTA TGGAAAACG CCAGCAACGC GGCCTTTTTA CGGTTCCCTG CCTTTTGCTG
6601 GCCTTTTGCT CACATGTTCT TTCCTGCGTT ATCCCCTGAT TCTGTGGATA ACCGTATTAC
6661 CGCCTTTGAG TGAGCTGATA CCGCTCGCCG CAGCCGAACG ACCGAGCGCA GCGAGTCAGT
6721 GAGCGAGGAA GCGGAAGAGC GCCCAATACG CAAACCGCCT CTCCCCGCGC GTTGGCCGAT
6781 TCATTAATGC AGCTGGCACG ACAGGTTTCC CGACTGGAAA GCGGGCAGTG AGCGCAACGC
6841 AATTAATGTG AGTTAGCTCA CTCATTAGGC ACCCCAGGCT TTACACTTTA TGCTTCCGGC
6901 TCGTATGTTG TGTGGAATTG TGAGCGGATA ACAATTTAC ACAGGAAACA GCTATGACCA
6961 TGATTACGCC AAGCGCGCAA TTAACCTCA CTAAAGGGAA CAAAAGCTGG GTACCGGGCC
7021 CCCCCCT

FIGURE 31D

76/240

Figure 32A: pDEST12.2 CMV Promoter for Eukaryotic Expression, SV40 Promoter/ori for G418 Resistance

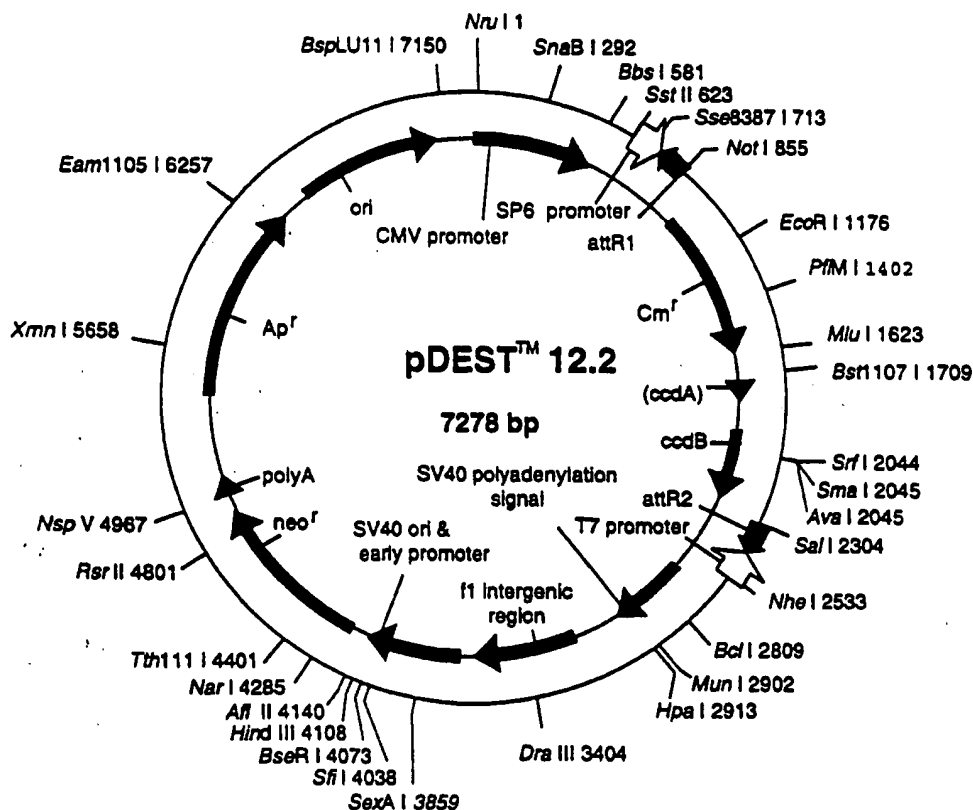
307 *mRNA from CMV promoter*
acc gtc aga tgc cct gga gac gcc atc cac gct gtt ttg acc tcc ata gaa
tgg cag tct agc gga cct ctg cgg tag gtg cga caa aac tgg agg tat ctt

358 gac acc ggg acc gat cca gcc tcc gga ctc tag cct agg ccg cgg agc gga
ctg tgg ccc tgg cta ggt cgg agg cct gag atc gga tcc ggc gcc tgc cct

409 taa caa ttt cac aca gga aac agc tat gac cat tag gcc ttt gca aaa agc
att gtt aaa gtg tgt cct ttg tgc ata ctg gta atc cgg aaa cgt ttt tgc

460 tat tta ggt gac act ata gaa ggt acg cct gca ggt acc ggt ccg gaa ttc
ata aat cca ctg tga tat ctt cca tgc gga cgt cca tgg cca ggc ctt aag
Int attR1

511 cca tca aca agt tgg taa ada gaa gct gaa cga gaa agc taa aat gat ata
ggt agt tgt tca aac atg ttt tta cga ctc gct ctt tgc att tta cca tat



77/240

pDEST12.2 7278 bp (rotated to position 3900)

<u>Location (Base Nos.)</u>		<u>Gene Encoded</u>
86..136		ori
220..742		CMV promoter
1059..935		attR1
1168..1827		CmR
1947..2031		inactivated ccdA
2169..2474		ccdB
2515..2639		attR2
2824..3186		small t & polyA
3310..3378		lac
4363..5157		neo
5680..6540		ampR

1	GGGGGGCGGA	GCCTATGGAA	AAACGCCAGC	AACGCGGCCCT	TTTTACGGTT	CCTGGCCTTT
61	TGCTGGCCTT	TTGCTCACAT	GTTCTTTCCT	GCGTTATCCC	CTGATTCTGT	GGATAACCGT
121	ATTACCGCCT	TTGAGTGAGC	TGATACCGCT	CGCCGCAGCC	GAACGACCGA	GCGCAGCGAG
181	TCAGTGAGCG	AGGAAGCGGA	AGAGCTCGCG	AATGCATGTC	GTTACATAAC	TTACGGTAAA
241	TGGCCCGCCT	GGCTGACCGC	CCAACGACCC	CCGCCCATTG	ACGTCAATAA	TGACGTATGT
301	TCCCATAGTA	ACGCCAATAG	GGACTTTCCA	TTGACGTCAA	TGGGTGGAGT	ATTACGGTA
361	AACTGCCCCAC	TTGGCAGTAC	ATCAAGTGTA	TCATATGCCA	AGTACGCCCC	CTATTGACGT
421	CAATGACGGT	AAATGGCCCC	CCTGGCATT	TGCCCAGTAC	ATGACCTTAT	GGGACTTTCC
481	TACTTGGCAG	TACATCTACG	TATTAGTCAT	CGCTATTACC	ATGGTGATGC	GGTTTGGCA
541	GTACATCAAT	GGGCGTGGAT	AGCGGTTTGA	CTCACGGGGA	TTTCCAAGTC	TCCACCCCAT
601	TGACGTCAAT	GGGAGTTTGT	TTTGGCACCA	AAATCAACGG	GACTTTCCAA	AATGTCGTAA
661	CAACTCCGCC	CCATTGACGC	AAATGGGCGG	TAGGCGTGTA	CGGTGGGAGG	TCTATATAAG
721	CAGAGCTCGT	TTAGTGAACC	GTCAGATCGC	CTGGAGACGC	CATCCACGCT	GTTTTGACCT
781	CCATAGAAGA	CACCGGGACC	GATCCAGCCT	CCGGA CTCTA	GCCTAGGCCG	CGGGACGGAT
841	AACAATTTCA	CACAGGAAAC	AGCTATGACC	ATTAGGCCTT	TGCAAAAAGC	TATTTAGGTG
901	ACACTATAGA	AGGTACGCCT	GCAGGTACCG	GATCACAAGT	TTGTACAAA	AAGCTGAACG
961	AGAAACGTAA	AATGATATAA	ATATCAATAT	ATTAAATTAG	ATTTTGCATA	AAAACAGAC
1021	TACATAATAC	TGTAAACAC	AACATATCCA	GTCATATGG	CGGCCGCATT	AGGCACCCCA
1081	GGCTTTACAC	TTTATGCTTC	CGGCTCGTAT	AATGTGTGGA	TTTTGAGTTA	GGATCCGTCG
1141	AGATTTTCAG	GAGCTAAGGA	AGCTAAAGTA	GAGAAAAAAA	TCACTGGATA	TACCACCGTT
1201	GATATATCCC	AATGGCATCG	TAAAGAACAT	TTTGAGGCAT	TTCACTCAGT	TGCTCAATGT
1261	ACCTATAACC	AGACCGTTCA	GCTGGATATT	ACGGCCTTTT	TAAAGACCGT	AAAGAAAAAT
1321	AAGCACAAGT	TTTATCCGGC	CTTTATTAC	ATTCTTGCCC	GCCTGATGAA	TGCTCATCCG
1381	GAATTCGGTA	TGGCAATGAA	AGACGGTGAG	CTGGTGATAT	GGGATAGTGT	TCACCTTGT
1441	TACACCGTTT	TCCATGAGCA	AACTGAAACG	TTTTTCATCGC	TCTGGAGTGA	ATACCACGAC
1501	GATTTCCGGC	AGTTTCTACA	CATATATTCTG	CAAGATGTGG	CGTGTACGG	TGAAAACCTG
1561	GCCTATTTCC	CTAAAGGGTT	TATTGAGAA	ATGTTTTTCG	TCTCAGCCAA	TCCCTGGGTG
1621	AGTTTTCACCA	GTTTTGATTT	AAACGTGGCC	AATATGGACA	ACTTCTTCGC	CCCCGTTTTTC
1681	ACCATGGGCA	AATATTATAC	GCAAGGCGAC	AAGGTGCTGA	TGCCGCTGGC	GATTACGGTT
1741	CATCATGCCG	TCTGTGATGG	CTTCCATGTC	GGCAGAATGC	TTAATGAATT	ACAACAGTAC
1801	TGCGATGAGT	GGCAGGGCGG	GGCGTAAACG	CGTGGATCCG	GCTTACTAAA	AGCCAGATAA
1861	CAGTATGCGT	ATTTGCGCGC	TGATTTTTTG	GGTATAAGAA	TATATACTGA	TATGTATACC
1921	CGAAGTATGT	CAAAAAGAGG	TGTGCTATGA	AGCAGCGTAT	TACAGTGACA	GTTGACAGCG
1981	ACAGCTATCA	GTTGCTCAAG	GCATATATGA	TGTCAATATC	TCCGGTCTGG	TAAGCACAAC
2041	CATGCAGAA	GAAGCCGTC	GTCGCGTGC	CGAACGCTGG	AAAGCGGAAA	ATCAGGAAGG
2101	GATGGCTGAG	GTCGCCCGGT	TTATTGAAAT	GAACGGCTCT	TTTGCTGACG	AGAACAGGGA
2161	CTGGTGAAAT	GCAGTTTAAG	GTTTACACCT	ATAAAAGAGA	GAGCCGTTAT	CGTCTGTTTG
2221	TGGATGTACA	GAGTGATATT	ATTGACACGC	CCGGGCGACG	GATGGTGATC	CCCCTGGCCA
2281	GTGCACGTCT	GCTGTGATAT	AAAGTCTCCC	GTGAACTTTA	CCCGGTGGTG	CATATCGGGG
2341	ATGAAAGCTG	GCGCATGATG	ACCAACCGATA	TGGCCAGTGT	GCCGGTCTCC	GTTATCGGGG
2401	AAGAAGTGCG	TGATCTCAGC	CACCGCGAAA	ATGACATCAA	AAACGCCATT	AACCTGATGT-

FIGURE 32B

78/240

2461 TCTGGGGAAT ATAAATGTCA GGCTCCCTTA TACACAGCCA GTCTGCAGGT CGACCATAGT
2521 GACTGGATAT GTTGTGTTTT ACAGTATTAT GTAGTCTGTT TTTTATGCAA AATCTAATTT
2581 AATATATTGA TATTTATATC ATTTTACGTT TCTCGTTCAG CTTTCTTGTA CAAAGTGGTG
2641 ATCGCGTGCA TGCACGTCA TAGCTCTCTC CCTATAGTGA GTCGTATTAT AAGCTAGGCA
2701 CTGGCCGTCG TTTTACAACG TCGTGACTGG GAAAACTGCT AGCTTGGGAT CTTTGTGAAG
2761 GAACCTTACT TCTGTGGTGT GACATAATTG GACAACTAC CTACAGAGAT TTAAAGCTCT
2821 AAGGTAAATA TAAAATTTTT AAGTGTATAA TGTGTTAAAC TAGCTGCATA TGCTTGCTGC
2881 TTGAGAGTTT TGCTTACTGA GTATGATTTA TGAAAATATT ATACACAGGA GCTAGTGAAT
2941 CTAATTGTTT GTGTATTTTA GATTCACAGT CCCAAGGCTC ATTTCAGGCC CCTCAGTCCT
3001 CACAGTCTGT TCATGATCAT AATCAGCCAT ACCACATTG TAGAGGTTTT ACTTGCTTTA
3061 AAAACCTCC CACACCTCCC CCTGAACCTG AAACATAAAA TGAATGCAAT TGTGTTGTT
3121 AACTTGTTTA TTGCAGCTTA TAATGGTTAC AAATAAAGCA ATAGCATCAC AAATTTTACA
3181 AATAAAGCAT TTTTTCCTACT GCATTCTAGT TGTGGTTTGT CCAAACCTCAT CAATGTATCT
3241 TATCATGTCT GGATCGATCC TGCATTAATG AATCGGCCAA CGCGCGGGGA GAGGCGGTTT
3301 GCGTATTGGC TGGCGTAATA GCGAAGAGGC CCGCACCGAT CGCCCTTCCC AACAGTTGCG
3361 CAGCTGTAAT GGCGAATGGG ACGCGCCCTG TAGCGGCGCA TTAAGCGCGG CGGGTGTGGT
3421 GGTTACGCGC AGCGTGACCG CTACACTTGC CAGCGCCCTA GCGCCCGCTC CTTTCGCTTT
3481 CTTCCCTTCC TTTCTCGCCA CGTTCGCGCG CTTTCCCGCT CAAGCTCTAA ATCGGGGGCT
3541 CCCTTTAGGG TTCCGATTTA GTGCTTTACG GCACCTCGAC CCCAAAAAAC TTGATTAGGG
3601 TGATGTTTCA CGTAGTGGGC CATCGCCCTG ATAGACGGTT TTTTCGCCCT TGACGTTGGA
3661 GTCCACGTTT TTTAATAGTG GACTCTTGTT CCAAACCTGGA ACAACACTCA ACCCTATCTC
3721 GGTCTATTCT TTTGATTTAT AAGGGATTTT GCCGATTTCT GCCTATTGGT TAAAAATGA
3781 GCTGATTTAA CAAATATTTA ACGCGAATTT TAACAAAATA TTAACGTTTA CAATTTTCGCC
3841 TGATGCGGTA TTTTCTCCTT ACGCATCTGT GCGGTATTTT ACACCGCATA CGCGGATCTG
3901 CGCAGCACCA TGGCCTGAAA TAACCTCTGA AAGAGGAACT TGGTTAGGTA CCTTCTGAGG
3961 CGGAAAGAAC CAGCTGTGGA ATGTGTGTCA GTTAGGGTGT GGAAAGTCCC CAGGCTCCCC
4021 AGCAGCGAGA AGTAGCAAAA GCATGCATCT CAATTAGTCA GCAACCAGGT GTGGAAAGTC
4081 CCCAGGCTCC CCAGCAGGCA GAAATATGCA AAGCATGCAT CTCAATTAGT CAGCAACCAT
4141 AGTCCCGCCC CTAACCTCCG CCATCCCGCC CCTAACCTCCG CCCAGTTCCG CCCATTCTCC
4201 GCCCCATGGC TGACTAATTT TTTTATTTA TGCAGAGGCC GAGGCCGCTT CGGCCTCTGA
4261 GCTATTCCAG AAGTAGTGAG GAGGCTTTTT TGGAGGCTTA GGCTTTTGCA AAAAGTTGA
4321 TTCTTCTGAC ACAACAGTCT CGAACTTAAG GCTAGAGCCA CCATGATTGA ACAAGATGGA
4381 TTGCACGCAG GTTCTCCGGC CGCTTGGGTG GAGAGGCTAT TCGGCTATGA CTGGGCACAA
4441 CAGACAATCG GCTGCTCTGA TGCCGCCGTG TTCCGGCTGT CAGCGCAGGG GCGCCCGGTT
4501 CTTTGTGTCA AGACCGACCT GTCCGGTGCC CTGAATGAAC TGCAGGACGA GGCAGCGCGG
4561 CTATCGTGCG TGCCACGAC GGGCGTTTCT TGCAGAGCTG TGCTCGACGT TGCTACTGAA
4621 GCGGGAAGGG ACTGGCTGCT ATTGGGCGAA GTGCCGGGGC AGGATCTCCT GTCATCTCAC
4681 CTTGCTCCTG CCGAGAAAGT ATCCATCATG GCTGATGCAA TGCGGCGGCT GCATACGCTT
4741 GATCCGGCTA CCGTCCCATT CGACCACCAA GCGAAACATC GCATCGAGCG AGCAGCTACT
4801 CGGATGGAAG CCGGTCTTGT CGATCAGGAT GATCTGGACG AAGAGCATCA GGGGCTCGCG
4861 CCAGCCGAAC TGTCGCCAG GCTCAAGGCG CGCATGCCCG ACGGCGAGGA TCTCGTCGTG
4921 ACCCATGGCG ATGCCTGCTT GCCGAATATC ATGGTGGAAG ATGGCCGCTT TTCTGGATT
4981 ATCGACTGTG GCCGGCTGGG TGTGGCGGAC CGCTATCAGG ACATAGCGTT GGCTACCCGT
5041 GATATTGCTG AAGAGCTTGG CGGCGAATGG GCTGACCGCT TCCTCGTGCT TTACGGTATC
5101 GCCGCTCCCC ATTCGCAGCG CATCGCCTTC TATCGCCTTC TTGACGAGTT CTTCTGAGCG
5161 GGACTCTGGG GTTCGAAATG ACCGACCAAG CGACGCCCAA CTGCCCATCA CGATGGCCCG
5221 AATAAAATAT CTTTATTTTC ATTACATCTG TGTGTTGGTT TTTTGTGTGA ATCGATAGCG
5281 ATAAGGATCC GCGTATGGTG CACTCTCAGT ACAATCTGCT CTGATGCCGC ATAGTTAAGC
5341 CAGCCCCGAC ACCCGCCAAC ACCCGCTGAC GCGCCCTGAC GGGCTTGTCT GCTCCCGGCA
5401 TCCGCTTACA GACAAGCTGT GACCGTCTCC GGGAGCTGCA TGTGTCAGAG GTTTTCACCG
5461 TCATCACCAG AACGCGCGAG ACGAAAGGGC CTCGTGATAC GCCTATTTTT ATAGGTTAAT
5521 GTCATGATAA TAATGGTTTC TTAGACGTCA GGTGGCACTT TTCGGGGAAG TGTGCGCGGA
5581 ACCCTTATTT GTTTATTTTT CTAAATACAT TCAAATATGT ATCCGCTCAT GAGACAATAA
5641 CCCTGATAAA TGCTTCAATA ATATTGAAAA AGGAAGAGTA TGAGTATTCA ACATTTCCGT
5701 GTCGCCCTTA TTCCCTTTTT TGCGGCATTT TGCCCTTCTG TTTTGTCTCA CCCAGAAACG
5761 CTGGTGAAAG TAAAAGATGC TGAAGATCAG TTGGGTGCAC GAGTGGGTTA CATCGAACTG
5821 GATCTCAACA GCGGTAAGAT CCTTGAGAGT TTTTCGCCCC AAGAAGCTTT TCCAATGATG
5881 AGCACTTTTA AAGTTCTGCT ATGTGGCGCG GTATTATCCC GTATTGACGC CGGGCAAGAG-

FIGURE 32C

79/240

5941 CAACTCGGTC GCCGCATACA CTATTCTCAG AATGACTTGG TTGAGTACTC ACCAGTCACA
6001 GAAAAGCATC TTACGGATGG CATGACAGTA AGAGAATTAT GCAGTGCTGC CATAACCATG
6061 AGTGATAACA CTGCGGCCAA CTTACTTCTG ACAACGATCG GAGGACCGAA GGAGCTAACC
6121 GCTTTTTTGC ACAACATGGG GGATCATGTA ACTCGCCTTG ATCGTTGGGA ACCGGAGCTG
6181 AATGAAGCCA TACCAAACGA CGAGCGTGAC ACCACGATGC CTGTAGCAAT GGCAACAACG
6241 TTGCGCAAAC TATTAACTGG CGAACTACTT ACTCTAGCTT CCCGGCAACA ATTAATAGAC
6301 TGGATGGAGG CGGATAAAGT TGCAGGACCA CTTCTGCGCT CGGCCCTTCC GGCTGGCTGG
6361 TTTATTGCTG ATAAATCTGG AGCCGGTGAG CGTGGGTCTC GCGGTATCAT TGCAGCACTG
6421 GGGCCAGATG GTAAGCCCTC CCGTATCGTA GTTATCTACA CGACGGGGAG TCAGGCAACT
6481 ATGGATGAAC GAAATAGACA GATCGCTGAG ATAGGTGCCT CACTGATTAA GCATTGGTAA
6541 CTGTCAGACC AAGTTTACTC ATATATACTT TAGATTGATT TAAACTTCA TTTTAAATTT
6601 AAAAGGATCT AGGTGAAGAT CCTTTTGTAT AATCTCATGA CCAAAATCCC TTAACGTGAG
6661 TTTTCGTTCC ACTGAGCGTC AGACCCCGTA GAAAAGATCA AAGGATCTTC TTGAGATCCT
6721 TTTTTTCTGC GCGTAATCTG CTGCTTGCAA ACAAAAAAAC CACCGCTACC AGCGGTGGTT
6781 TGTTTGCCGG ATCAAGAGCT ACCAACTCTT TTTCCGAAGG TAACTGGCTT CAGCAGAGCG
6841 CAGATACCAA ATACTGTCCT TCTAGTGTAG CCGTAGTTAG GCCACCACTT CAAGAACTCT
6901 GTAGCACCGC CTACATACCT CGTCTGCTA ATCCTGTTAC CAGTGGCTGC TGCCAGTGGC
6961 GATAAGTCGT GTCTTACCGG GTTGGACTCA AGACGATAGT TACCGGATAA GGCGCAGCGG
7021 TCGGGCTGAA CGGGGGGTTC GTGCACACAG CCCAGCTTGG AGCGAACGAC CTACACCGAA
7081 CTGAGATACC TACAGCGTGA GCATTGAGAA AGCGCCACGC TTCCCGAAGG GAGAAAGGCG
7141 GACAGGTATC CGGTAAGCGG CAGGTCGGA ACAGGAGAGC GCACGAGGGA GCTTCCAGGG
7201 GGAAACGCCT GGTATCTTTA TAGTCCTGTC GGGTTTCGCC ACCTCTGACT TGAGCGTCGA
7261 TTTTGTGAT GCTCGTCA

FIGURE 32D

80/240

Figure 33A1

pDEST13

Native protein in E. coli: λ PL
promoter

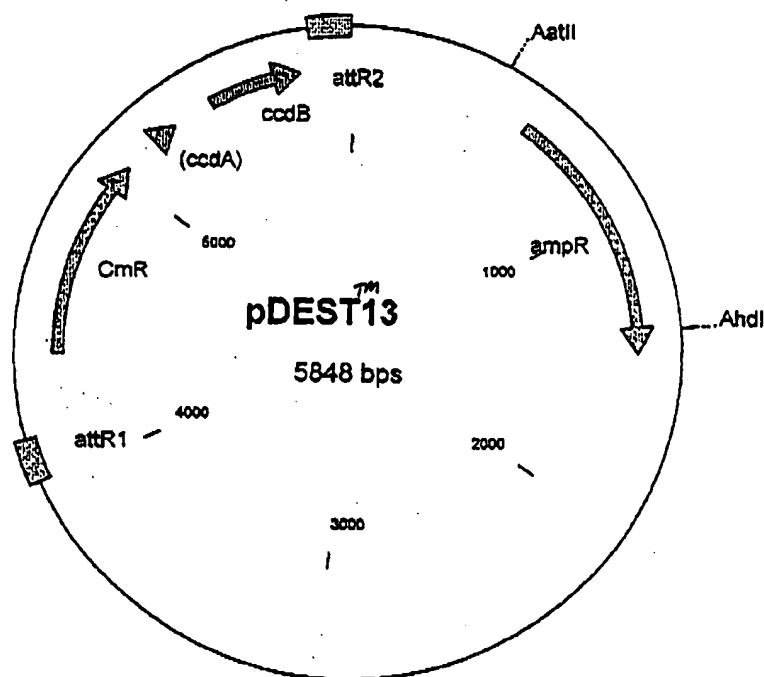
3721 tgggcaaacc aagacagcta aagatctctc acctaccaaa caatgcccc ctgcaaaaaa
 acccgtttgg ttctgtcgat ttctagagag tggatggttt gttacggggg gacgtttttt

3781 taaattcata taaaaaacat acagataacc atctgcggtg ataaattatc tctggcggtg
 atttaagtat attttttgta tgtctattgg tagacgccac tatttaatag agaccgccac

3841 ttgacataaa taccactggc ggtgatactg agcacatcag caggacgcac tgaccaccat
aactgtattt atgggtgaccg ccactatgac tctgttagtc gtctgcgtg actggtggta

3901 gaaggtgacg ctcttaaaaa ttaagecctg aagaaggga gcattcaaag cagaaggctt
 cttccactgc gagaattttt aattcgggac tcttcccgt cgtaagtttc gtcttcgaa

3961 tggggtgtgt gatacgaac gaagcattgg gatcatcaca agtttgtaga aaaaagctga
 accccacaca ctatgctttg cttcgtaacc ctagtatgt tcaaacatgt ttttcgact



81/240

pDEST13 5848 bp

<u>Location (Base Nos.)</u>		<u>Gene Encoded</u>
599..1458		ampR
4123..3998		attR1
4372..5031		CmR
5151..5235		inactivated ccdA
5373..5678		ccdB
5719..5843		attR2
1 TTCACTGGCC GTCGTTTTAC AACGTCGTGA CTGGGAAAAC CCTGGCGTTA CCCAACTTAA		
61 TCGCCTTGCA GCACATCCCC CTTTCGCCAG CTGGCGTAAT AGCGAAGAGG CCCGCACCGA		
121 TCGCCCTTCC CAACAGTTGC GCAGCCTGAA TGGCGAATGG CGCCTGATGC GGTATTTTCT		
181 CCTTACGCAT CTGTGCGGTA TTTCACACCG CATATGGTGC ACTCTCAGTA CAATCTGCCTC		
241 TGATGCCGCA TAGTTAAGCC AGCCCCGACA CCCGCCAACA CCCGCTGACG CGCCCTGACG		
301 GGCTTGTCTG CTCCCGGCAT CCGCTTACAG ACAAGCTGTG ACCGTCTCCG GGAGCTGCAT		
361 GTGTCAGAGG TTTTCACCGT CATCACCGAA ACGCGCGAGA CGAAAGGGCC TCGTGATACG		
421 CCTATTTTAA TAGGTTAATG TCATGATAAT AATGGTTTCT TAGACGTCAG GTGGCACTTT		
481 TCGGGGAAAT GTGCGCGGAA CCCCTATTTC TTTATTTTTC TAAATACATT CAAATATGTA		
541 TCCGCTCATG AGACAATAAC CCTGATAAAT GCTTCAATAA TATTGAAAAA GGAAGAGTAT		
601 GAGTATTCAA CATTTCCGTG TCGCCCTTAT TCCCTTTTTC GCGGCATTTT GCCTTCCTGT		
661 TTTTGCTCAC CCAGAAACGC TGGTGAAAGT AAAAGATGCT GAAGATCAGT TGGGTGCACG		
721 AGTGGGTTAC ATCGAACTGG ATCTCAACAG CGGTAAGATC CTTGAGAGTT TTCGCCCCGA		
781 AGAACGTTTT CCAATGATGA GCACTTTTAA AGTTCTGCTA TGTGGCGCGG TATTATCCCG		
841 TATTGACGCC GGGCAAGAGC AACTCGGTGCG CCGCATACAC TATTCTCAGA ATGACTTGGT		
901 TGAGTACTCA CCAGTCACAG AAAAGCATCT TACGGATGGC ATGACAGTAA GAGAATTATG		
961 CAGTGCTGCC ATAACCATGA GTGATAACAC TGCGGCCAAC TTAATTCTGA CAACGATCGG		
1021 AGGACCGAAG GAGCTAACCG CTTTTTTGCA CAACATGGGG GATCATGTAA CTCGCCTTGA		
1081 TCGTTGGGAA CCGGAGCTGA ATGAAGCCAT ACCAAACGAC GAGCGTGACA CCACGATGCC		
1141 TGTAGCAATG GCAACAACGT TCGCCAAACCT ATTAAGTGGC GAACTACTTA CTCTAGCTTC		
1201 CCGGCAACAA TTAATAGACT GGATGGAGGC GGATAAAGTT GCAGGACCAC TTCTGCGCTC		
1261 GGCCCTTCCG GCTGGCTGGT TTATTGCTGA TAAATCTGGA GCCGGTGAGC GTGGGTCTCG		
1321 CGGTATCATT GCAGCACTGG GGCCAGATGG TAAGCCCTCC CGTATCGTAG TTATCTACAC		
1381 GACGGGGAGT CAGGCAACTA TGGATGAACG AAATAGACAG ATCGCTGAGA TAGGTGCTTC		
1441 ACTGATTAAG CATTGGTAAC TGTCAGACCA AGTTTACTCA TATATACTTT AGATTGATTT		
1501 AAAAATTTCAT TTTTAATTTA AAAGGATCTA GGTGAAGATC CTTTTTGATA ATCTCATGAC		
1561 CAAAATCCCT TAACGTGAGT TTTTCGTTCCA CTGAGCGTCA GACCCCGTAG AAAAGATCAA		
1621 AGGATCTTCT TGAGATCCTT TTTTCTGCG CGTAATCTGC TGCTTGCAAA CAAAAAACC		
1681 ACCGCTACCA GCGGTGGTTT GTTTGCCGGA TCAAGAGCTA CCAACTCTTT TTCCGAAGGT		
1741 AACTGGCTTC AGCAGAGCGC AGATAACAAA TACTGTTCTT CTAGTGATAGC CGTAGTTAGG		
1801 CCACCACTTC AAGAACTCTG TAGCACCGCC TACATACCTC GCTCTGCTAA TCCTGTTACC		
1861 AGTGGCTGCT GCCAGTGGCG ATAAGTCGTG TCTTACCGGG TTGGACTCAA GACGATAGTT		
1921 ACCGGATAAG GCGCAGCGGT CCGGCTGAAC GGGGGGTTTC TGCACACAGC CCAGCTTGGA		
1981 GCGAACGACC TACACCGAAC TGAGATACCT ACAGCGTGAG CATTGAGAAA GCGCCACGCT		
2041 TCCCGAAGGG AGAAAGGCGG ACAGGTATCC GGTAAGCGGC AGGGTCGGAA CAGGAGAGCG		
2101 CACGAGGGAG CTTCAGGGG GAAACGCTG GTATCTTTAT AGTCCTGTCG GGTTCGCCA		
2161 CCTCTGACTT GAGCGTGGAT TTTTGTGATG CTCGTACAGG GGGCGGAGCC TATGGAAAAA		
2221 CGCCAGCAAC GCGGCCCTTT TACGGTTTCT GGCCTTTTTC TGGCCTTTTG CTCACATGTT		
2281 CTTTCTGCG TTATCCCTTG ATTCTGTGGA TAACCGTATT ACCGCCTTTG AGTGAGCTGA		
2341 TACCGTCTGC CGCAGCCGAA CGACCGAGCG CAGCGAGTCA GTGAGCGAGG AAGCGGAAGA		
2401 GCGCCCAATA CGCAAACCGC CTCTCCCCGC GCGTTGGCCG ATTCATTAAT GCAGCTGGCA		
2461 CGACAGGTTT CCCGACTGGA AAGCGGGCAG TGAGCGCAAC GCAATTAATG TGAGTTAGCT		
2521 CACTCATTAG GCACCCAGG CTTTACACTT TATGCTTCCG GCTCGTATGT TGTGTGGAAT		
2581 TGTGAGCGGA TAACAATTTT ACACAGGAAA CAGCTATGAC CATGATTACG CCAAGCTTGG		
2641 CTGCAGGTGA TGATTATCAG CCAGCAGAGA TTAAGGAAAA CAGACAGGTT TATTGAGCGC		
2701 TTATCTTTCC CTTTATTTT GTCGCGGTAA GTCGCATAAA AACCATTCTT CATAATTCAA-		

FIGURE 33B

82/240

2761 TCCATTTACT ATGTTATGTT CTGAGGGGAG TGAATTTCC CTAATTCGA TGAAGATTCT
2821 TGCTCAATTG TTATCAGCTA TGCGCCGACC AGAACACCTT GCCGATCAGC CAAACGTCTC
2881 TTCAGGCCAC TGACTAGCGA TAACCTTTCCC CACAACGGAA CAACTCTCAT TGCATGGGAT
2941 CATTGGGGTAC TGTGGGTTTA GTGGTTGTAA AAACACCTGA CCGCTATCCC TGATCAGTTT
3001 CTGAAGGTA AACTCATCAC CCCCAGTCT GGCTATGCAG AAATCACCTG GCTCAACAGC
3061 CTGCTCAGGG TCAACGAGAA TTAACATTCC GTCAGGAAAG CTGGCTTGG AGCCTGTTGG
3121 TGCGGTCATG GAATTACCTT CAACCTCAAG CCAGAATGCA GAATCACTGG CTTTTTTGGT
3181 TGTGCTTACC CATCTCTCCG CATCACCTTT GGTAAAGGTT CTAAGCTTAG GTGAGAACAT
3241 CCCTGCCTGA ACATGAGAAA AAACAGGGTA CTCATACTCA CTTCTAAGTG ACGGCTGCAT
3301 ACTAACCGCT TCATACATCT CGTAGATTTC TCTGGCGATT GAAGGGCTAA ATTCTTCAAC
3361 GCTAACTTTG AGAATTTTTG CAAGCAATGC GCGGTTATAA GCATTTAATG CATTGATGCC
3421 ATTAATAAAA GCACCAACGC CTGACTGCCC CATCCCCATC TTGTCTGCGA CAGATTCCTG
3481 GGATAAGCCA AGTTCATTTT TCTTTTTTTT TGTGCTCATA CGTTAAATCT ATCACCAGCA
3541 AAGCTGCTCT TGTGTTAATG GTTCTTTTTT TGTGCTCATA CGTTAAATCT ATCACCAGCA
3601 GGGATAAATA TCTAACACCG TCGTGTGTA CTATTTTACC TCTGGCGGTG ATAATGGTTG
3661 CATGTACTAA GGAGGTTGTA TGAACAACG CATAACCCTG AAAGATTATG CAATGCGCTT
3721 TGGGCAAACC AAGACAGCTA AAGATCTCTC ACCTACCAA CAATGCCCC CTGCAAAAAA
3781 TAAATTCATA TAAAAACAT ACAGATAACC ATCTGCGGTG ATAAATTATC TCTGGCGGTG
3841 TTGACATAAA TACCACTGGC GGTGATACTG AGCACATCAG CAGGACGCAC TGACCACCAT
3901 GAAGGTGACG CTCTTAAAAA TTAAGCCCTG AAGAAGGGCA GCATTCAAAG CAGAAGGCTT
3961 TGGGGTGTGT GATACGAAAC GAAGCATTGG GATCATCACA AGTTTGTACA AAAAAGCTGA
4021 ACGAGAAACG TAAATGATA TAAATATCAA TATATTAAAT TAGATTTTGC ATAAAAACA
4081 GACTACATAA TACTGTAAAA CACAACATAT CCAGTCACTA TGGCGGCCGC TAAGTTGGCA
4141 GCATCACCAG ACGCACTTTG CGCCGAATAA ATACCTGTGA CGGAAGATCA CTTCCGAGAA
4201 TAAATAAATC CTGGTGTCCC TGTTGATACC GGGGAAGCCCT GGGCCAACTT TTGGCGAAAA
4261 TGAGACGTTG ATCGGCACGT AAGAGGTTCC AACTTTCACC ATAATGAAAT AAGATCACTA
4321 CCGGGCGTAT TTTTGTAGTT ATCGAGATT TACAGAGCTA AGGAAGCTAA AATGGAGAAA
4381 AAAATCACTG GATATACCAC CGTTGATATA TCCCAATGGC ATCGTAAAGA ACATTTTGAG
4441 GCATTTTCAGT CAGTTGTCTA ATGTACCTAT AACCAGACCG TTCAGCTGGA TATTACGGCC
4501 TTTTAAAGGA CCGTAAAGAA AAATAAGCAC AAGTTTTATC CGGCCTTTAT TCACATTCTT
4561 GCGCGCTGA TGAATGCTCA TCCGGAATTC CGTATGGCAA TGAAGACGG TGAGCTGGTG
4621 ATATGGGATA GTGTTCAACC TTGTTACACC GTTTTCCATG AGCAAACCTGA AACGTTTTCA
4681 TCGCTCTGGA GTGAATACCA CGACGATTTT CCGCAGTTTC TACACATATA TTCGCAAGAT
4741 GTGGCGTGT ACGGTGAAAA CCTGGCCTAT TTCCCTAAAG GGTATTATGA GAATATGTTT
4801 TCGTCTCAG CCAATCCCCTG GGTGAGTTTC ACCAGTTTGG ATTTAAACGT GGCCAATATG
4861 GACAACTTCT TCGCCCCCGT TTTCAACATG GGCAAATATT ATACGCAAGG CGACAAGGTG
4921 CTGATGCCGC TGGCGATTCA GGTTCATCAT GCCGTCTGTG ATGGCTTCCA TGTCCGCAGA
4981 ATGCTTAATG AATTACAACA GTACTGCGAT GAGTGGCAGG GCGGGGCGTA AACCGTGGA
5041 TCCGGCTTAC TAAAAGCCAG ATAACAGTAT GCGTATTTGC GCGCTGATTT TTGCGGTATA
5101 AGAATATATA CTGATATGTA TACCCGAAGT ATGTCAAAAA GAGGTGTGCT ATGAAGCAGC
5161 GTATTACAGT GACAGTTGAC AGCGACAGCT ATCAGTTGCT CAAGGCATAT ATGATGTCAA
5221 TATCTCCGGT CTGGTAAGCA CAACATGCA GAATGAAGCC CGTCGTCTGC GTGCCGAACG
5281 CTGGAAAGCG GAAAATCAGG AAGGGATGGC TGAGGTCGCC CGGTTTATTG AAATGAACGG
5341 CTCTTTTGCT GACGAGAACA GGGACTGGTG AAATGCAGTT TAAGGTTTAC ACCTATAAAA
5401 GAGAGAGCCG TTATCGTCTG TTTGTGGATG TACAGAGTGA TATTATTGAC ACGCCCGGGC
5461 GACGGATGGT GATCCCCCTG GCCAGTGCAC GTCTGCTGTC AGATAAAGTC TCCCGTGAAC
5521 TTTACCCGGT GGTGCATATC GGGGATGAAA GCTGGCGCAT GATGACCACC GATATGGCCA
5581 GTGTGCCGGT CTCGTTATC GGGGAAGAAG TGGCTGATCT CAGCCACCGC GAAAATGACA
5641 TCAAAAACGC CATTAACCTG ATGTTCTGGG GAATATAAAT GTCAGGCTCC GTTATACACA
5701 GCCAGTCTGC AGGTGCACCA TAGTGACTGG ATATGTTGTG TTTTACAGTA TTATGTAGTC
5761 TGTTTTTTAT GCAAAATCTA ATTTAATATA TTGATATTTA TATCATTTTA CGTTTCTCGT
5821 TCAGCTTTCT TGTACAAAGT GGTGATAA

FIGURE 33C

84/240

pDEST14 6422 bp (rotated to position 4000)

<u>Location (Base Nos.)</u>		<u>Gene Encoded</u>
185..61		attR1
435..1094		CmR
1214..1298		inactivated ccdA
1436..1741		ccdB
1782..1906		attR2
2632..3489		ampR
1	CGATCCCGCG AAATTAATAC GACTCACTAT AGGGAGACCA CAACGGTTTC CCTCTAGATC	
61	ACAAGTTTGT ACAAAAAAGC TGAACGAGAA ACGTAAATG ATATAAATAT CAATATATTA	
121	AATTAGATTT TGCATAAAAA ACAGACTACA TAATACTGTA AAACACAACA TATCCAGTCA	
181	CTATGGCGGC CGCTAAGTTG GCAGCATCAC CCGACGCACT TTGCGCCGAA TAAATACCTG	
241	TGACGGAAGA TCACTTCGCA GAATAAATAA ATCCTGGTGT CCCTGTTGAT ACCGGGAAGC	
301	CCTGGGCCAA CTTTGGCGA AAATGAGACG TTGATCGGCA CGTAAGAGGT TCCAACTTTC	
361	ACCATAATGA AATAAGATCA CTACCGGGCG TATTTTTTGA GTTATCGAGA TTTTCAGGAG	
421	CTAAGGAAGC TAAATGGAG AAAAAATCA CTGGATATAC CACCGTTGAT ATATCCCAAT	
481	GGCATCGTAA AGAACATTTT GAGGCATTTT AGTCAGTTGC TCAATGTACC TATAACCAGA	
541	CCGTTACGCT GGATATTACG GCCTTTTTAA AGACCGTAAA GAAAAATAAG CACAAGTTTT	
601	ATCCGGCCCT TATTACATT CTTGCCGCC TGATGAATGC TCATCCGGAA TTCCGTATGG	
661	CAATGAAAGA CGGTGAGCTG GTGATATGGG ATAGTGTTCA CCCTTGTTAC ACCGTTTTCC	
721	ATGAGCAAAC TGAAACGTTT TCATCGCTCT GGAGTGAATA CCACGACGAT TTCCGGCAGT	
781	TTCTACACAT ATATTCGCAA GATGTGGCGT GTTACGGTGA AAACCTGGCC TATTTCCCTA	
841	AAGGGTTTAT TGAGAATATG TTTTTCGTCT CAGCCAATCC CTGGGTGAGT TTCACCAGTT	
901	TTGATTTTAA CGTGGCCAAT ATGGACAAT TCTTCGCCCC CGTTTTCCACC ATGGGCAAAAT	
961	ATTATAGGCT AGGCGACAAG GTGCTGATGC CGCTGGCGAT TCAGGTTTAT CATGCCGTCT	
1021	GTGATGGCTT CCATGTCGCG AGAATGCTTA ATGAATTACA ACAGTACTGC GATGATGGC	
1081	AGGGCGGGGC GTAAACGCGT GGATCCGGCT TACTAAAAGC CAGATAACAG TATGCGTATT	
1141	TGCGCGCTGA TTTTTCGGT ATAAGAATAT ATACTGATAT GTATACCCGA AGTATGTCAA	
1201	AAAGAGGTGT GCTATGAAGC AGCGTATTAC AGTGACAGT GACAGCGACA GCTATCAGTT	
1261	GCTCAAGGCA TATATGATGT CAATATCTCC GGTCTGGTAA GCACAACCAT GCAGAAATGAA	
1321	GCCCGTCGTC TGCGTGCCGA ACGCTGGAAG GCGGAAATC AGGAAGGGAT GGCTGAGGTC	
1381	GCCCGGTTTA TTGAAATGAA CGGCTCTTTT GCTGACGAGA ACAGGGACTG GTGAAATGCA	
1441	GTTTAAGGTT TACACCTATA AAAGAGAGAG CCGTTATCGT CTGTTTGTGG ATGTACAGAG	
1501	TGATATTATT GACACGCCCG GGGCAGCGAT GGTGATCCCC CTGGCCAGTG CACGTCGTCT	
1561	GTCAGATAAA GTCTCCCGTG AACTTTACCC GGTGGTGCAT ATCGGGGATG AAAGCTGGCG	
1621	CATGATGACC ACCGATATGG CCAGTGTGCC GGTCTCCGTT ATCGGGGAAG AAGTGGCTGA	
1681	TCTCAGCCAC CGCGAAAATG ACATCAAAAA CGCCATTAACT CTGATGTTCT GGGGAATATA	
1741	AATGTCAAGC TCCCTTATAC ACAGCCAGTC TGCAGGTCGA CCATAGTGAC TGGATATGTT	
1801	GTGTTTTTACA GTATTATGTA GTCTGTTTTT TATGCAAAAT CTAATTTAAT ATATTGATAT	
1861	TTATATCATT TTACGTTTCT CGTTCAGCTT TCTTGTAACA AGTGGTGATG ATCCGGCTGC	
1921	TAACAAAGCC CGAAAAGGAG CTGAGTTGGC TGCTGCCACC GCTGAGCAAT AACTAGCATA	
1981	ACCCCTTGGG GCCTCTAAAC GGGTCTTGAG GGGTTTTTTG CTGAAAGGAG GAACTATATC	
2041	CGGATATCCA CAGGACGGGT GTGGTCGCCA TGATCGCGTA GTCGATAGTG GCTCCAAGTA	
2101	GCGAAGCGAG CAGGACTGGG CGGCGGCCAA AGCGGTCGGA CAGTGCTCCG AGAACGGGTG	
2161	CGCATAGAAA TTGCATCAAC GCATATAGCG CTAGCAGCAC GCCATAGTGA CTGGCGATGC	
2221	TGTCGGAATG GACGATATCC CGCAAGAGGC CCGGCAGTAC CGGCATAACC AAGCCTATGC	
2281	CTACAGCATC CAGGGTGACG GTGCCGAGGA TGACGATGAG CGCATTGTTA GATTTTCATAC	
2341	ACGGTGCCCTG ACTGCGTTAG CAATTAACT GTGATAAACT ACCGCATTA AGCTTATCGA	
2401	TGATAAGCTG TCAAACATGA GAATTCCTGA AGACGAAAGG GCCTCGTGAT ACGCCTATTT	
2461	TTATAGGTTA ATGTCATGAT AATAATGGTT TCTTAGACGT CAGGTGGCAC TTTTCGGGGA	
2521	AATGTGCGCG GAACCCCTAT TTGTTTATTT TTCTAAATAC ATTCAAATAT GTATCCGCTC	
2581	ATGAGACAAT AACCCTGATA AATGCTTCAA TAATATTGAA AAAGGAAGAG TATGAGTATT	
2641	CAACATTTCC GTGTCGCCCT TATTCCTTTT TTGCGGCAT TTTGCCTTCC TGTTTTGTCT	
2701	CACCCAGAAA CGCTGGTGAA AGTAAAAGAT GCTGAAGATC AGTTGGGTGC ACGAGTGGGT-	

FIGURE 34B


```

2761 TACATCGAAC TGGATCTCAA CAGCGGTAAG ATCCTTGAGA GTTTTCGCCC CGAAGAACGT
2821 TTTCGAATGA TGAGCACTTT TAAAGTTCTG CTATGTGGCG CCGTATTATC CCGTGTGAC
2881 GCCGGGCAAG AGCAACTCGG TCGCCGCATA CACTATTCTC AGAATGACTT GGTTGAGTAC
2941 TCACCAGTCA CAGAAAAGCA TCTTACGGAT GGCATGACAG TAAGAGAATT ATGCAGTGCT
3001 GCCATAACCA TGAGTGATAA CACTGCGGCC AACTTACTTC TGACAACGAT CGGAGGACCG
3061 AAGGAGCTAA CCGCTTTTTT GCACAAATG GGGGATCATG TAACTCGCCT TGATCGTTGG
3121 GAACCGGAGC TGAATGAAGC CATACCAAAC GACGAGCGTG ACACCACGAT GCCTGCAGCA
3181 ATGGCAACAA CGTTGCGCAA ACTATTAAC TGGCGAACTAC TTAATCTAGC TTCCCGGCAA
3241 CAATTAATAG ACTGGATGGA GCGGATAAAA GTTGACGAGC CACTTCTGCG CTCGGCCCTT
3301 CCGGCTGGCT GGTATTATGC TGATAAATCT GGAGCCGGTG AGCGTGGGTC TCGCGGTATC
3361 ATTGCAGCAC TGGGGCCAGA TGGTAAGCCC TCCCGTATCG TAGTTATCTA CACGACGGGG
3421 AGTCAGGCAA CTATGGATGA ACGAAATAGA CAGATCGCTG AGATAGGTGC CTCAGTGATT
3481 AAGCATTTGAT AACTGTCAGA CCAAGTTTAC TCATATATAC TTAGATTGA TTTAAACTT
3541 CATTTTAAAT TTAAAAGGAT CTAGGTGAAG ATCCTTTTGT ATAATCTCAT GACCAAAATC
3601 CCTTAACGTG AGTTTTCGTT CCACTGAGCG TCAGACCCCG TAGAAAAGAT CAAAGGATCT
3661 TCTTGAGATC CTTTTTTTCT GCGCGTAATC TGCTGCTTGC AAACAAAAAA ACCACCGCTA
3721 CCAGCGGTGG TTTGTTTGCC GGATCAAGAG CTACCAACTC TTTTCCGAA GGTAAGTGGC
3781 TTCAGCAGAG CGCAGATACC AAATACTGTC CTTCTAGTGT AGCCGTAGTT AGGCCACCAC
3841 TTCAAGAACT CTGTAGCACC GCCTACATAC CTCGCTCTGC TAATCCTGTT ACCAGTGGCT
3901 GCTGCCAGTG GCGATAAGTC GTGTCTTACC GGGTTGGACT CAAGACGATA GTTACGGAT
3961 AAGGCGCAGC GGTGCGGCTG AACGGGGGTT TCGTGACAC AGCCAGCTT GGAGCGAACG
4021 AACCTACACG AACTGAGATA CCTACAGCGT GAGCTATGAG AAAGCGCCAC GCTTCCGAA
4081 GGGAGAAAGG CGGACAGGTA TCCGGTAAGC GGCAGGGTCG GAACAGGAGA GCGCACGAGG
4141 GAGCTTCCAG GGGGAAACGC CTGGTATCTT TATAGTCCTG TCGGGTTTCG CCACCTCTGA
4201 CTTGAGCGTC GATTTTGTG ATGCTCGTCA GGGGGGCGGA GCCTATGGAA AAACGCCAGC
4261 AACGCGGCCT TTTTACGGTT CCTGGCCTTT TGCTGGCCTT TTGCTCATAT GTTCTTTCCT
4321 GCGTTATCCC CTGATTCTGT GGATAACCGT ATTACCGCCT TTGAGTGAGC TGATACCGCT
4381 CGCCGACGCG GAACGACCGA GCGCAGCGAG TCAGTGAGCG AGGAAGCGGA AGAGCGCCTG
4441 ATGCGGTATT TTCTCCTTAC GCATCTGTGC GGTATTTTAC ACCGCATATA TGGTGCACTC
4501 TCAGTACAAT CTGCTCTGAT GCCGCATAGT TAAGCCAGTA TAACTCCGC TATCGCTACG
4561 TGACTGGGTC ATGGCTGCGC CCCGACACCC GCCAACACCC GCTGACGCGC CCTGACGGGC
4621 TTGTCTGCTC CCGGCATCCG CTTACAGACA AGCTGTGACC GTCTCCGGGA GCTGCATGTG
4681 TCAGAGGTTT TCACCGTCAT CACCGAAACG CGCGAGGCAG CTGCGGTAAA GCTCATCAGC
4741 GTGGTCGTGA AGCGATTAC AGATGTCTGC CTGTTTATCC GCGTCCAGCT CGTTGAGTTT
4801 CTCCAGAAGC GTTAATGTCT GGCTTCTGAT AAAGCGGGCC ATGTTAAGGG CGGTTTTCCT
4861 ATGTTTTGTC ACTGATGCCT CCGTGAAGG GGGATTTCTG TTCATGGGGG TAATGATACC
4921 GATGAAACGA GAGAGGATGC TCACGATACG GGTACTGAT GATGAACATG CCCGGTACT
4981 GGAACGTTGT GAGGGTAAAC AACTGGCGGT ATGGATGCGG CGGGACCAGA GAAAAATCAC
5041 TCAGGGTCAA TGCCAGCGCT TCGTTAATAC AGATGTAGGT GTTCCACAGG GTAGCCAGCA
5101 GCATCCTCGC ATGCAGATCC GGAACATAAT GGTGCAGGGC GCTGACTTCC GCGTTTCCAG
5161 ACTTTACGAA ACACGGAAAC CGAAGACCAT TCATGTTGTT GCTCAGGTCG CAGACGTTTT
5221 GCAGCAGCAG TCGCTTCACG TTCGCTCGCG TATCGGTGAT TCATTCTGCT AACCAGTAAG
5281 GCAACCCCGC CAGCCTAGCC GGGTCCTCAA CGACAGGAGC ACGATCATGC GCACCCGTGG
5341 CCAGGACCCA ACGCTGCCCC AGATGCGCCG CGTGCGGCTG CTGGAGATGG CGGACGCGAT
5401 GGATATGTTT TGCCAAGGGT TGGTTTGCGC ATTACAGTCT CCGCAAGA ATTGATTGGC
5461 TCCAATTCTT GGAGTGGTGA ATCCGTTAGC GAGGTGCCGC CGGCTTCCAT TCAGGTGAGG
5521 GTGGCCCGGC TCCATGCACC GCGACGCAAC GCGGGGAGGC AGACAAGGTA TAGGGCGGCG
5581 CCTACAATCC ATGCCAACCC GTTCCATGTG CTCGCCGAGG CGGCATAAAT CGCCGTGACG
5641 ATCAGCGGTC CAGTGATCGA AGTTAGGCTG GTAAGAGCCG CGAGCGATCC TTGAAGCTGT
5701 CCGTGATGGT CGTCATCTAC CTGCCTGGAC AGCATGGCCT GCAACGCGGG CATCCCATG
5761 CCGCCGGAAG CGAGAAGAAT CATAATGGGG AAGGCCATCC AGCCTCGCGT CGCGAAGCCG
5821 AGCAAGACGT AGCCCAGCGC GTCGGCCGCC ATGCCGGCGA TAATGGCCTG CTTCTCGCCG
5881 AAACGTTTGG TGGCGGGACC AGTGACGAAG GCTTGAGCGA GGGCGTGCAA GATTCCGAAT
5941 ACCGCAAGCG ACAGGCGGAT CATCGTCGCG CTCCAGCGAA AGCGGTCTCT GCCGAAAATG
6001 ACCCAGAGCG CTGCCGCGAC CTGTCTACG AGTTGCATGA TAAAGAAGAC AGTCATAAGT
6061 GCGGCGACGA TAGTCATGCC CCGCGCCAC CGGAAGGAGC TGACTGGGTT GAAGGCTCTC
6121 AAGGGCATCG GTCGATCGAC GCTCTCCCTT ATGCGACTCC TGCATTAGGA AGCAGCCAG
6181 TAGTAGGTTG AGGCCGTTGA GCACGCGCGC CGCAAGGAAT GGTGCATGCA AGGAGATGGC

```

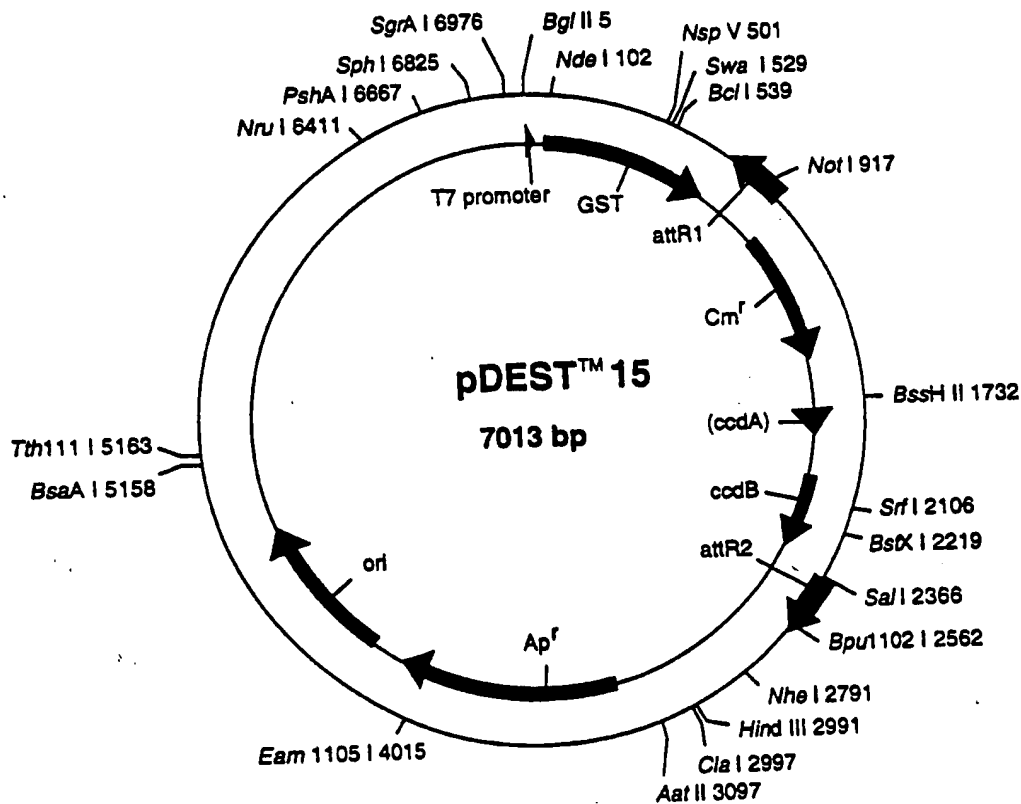
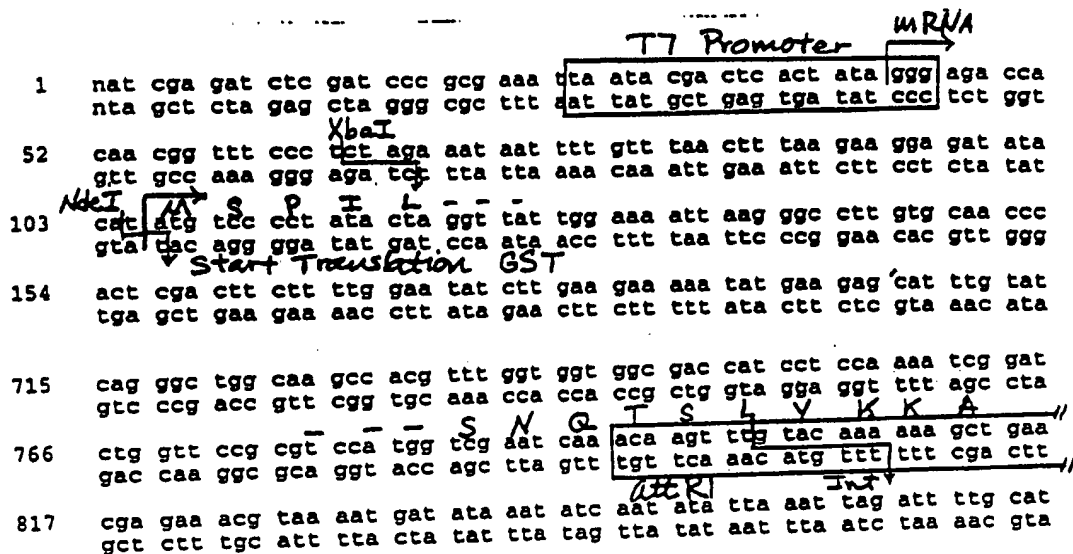
FIGURE 34C

86/240

6241 GCCCAACAGT CCCCCGGCCA CGGGGCCTGC CACCATACCC ACGCCGAAAC AAGCGTCAT
6301 GAGCCCGAAG TGGCGAGCCC GATCTTCCCC ATCGGTGATG TCGGCGATAT AGGCGCCAGC
6361 AACC GCACCT GTGGCGCCGG TGATGCCGGC CACGATGCGT CCGGCGTAGA GGATCGAGAT
6421 CT

FIGURE 34D

Figure 35A: pDEST15 Glutathione-S-transferase Fusion in *E. coli*, T7 Promoter



88/240

pDEST15 7013 bp

<u>Location (Base Nos.)</u>		<u>Gene Encoded</u>
108..776		GST
916..792		attR1
1025..1537		CmR
1804..1888		inactivated ccdA
2026..2331		ccdB
2372..2496		attR2
3233..4093		ampR

1	ATCGAGATCT	CGATCCCGCG	AAATTAATAC	GACTCACTAT	AGGGAGACCA	CAACGGTTTC
61	CCTCTAGAAA	TAATTTTGTT	TAACTTTAAG	AAGGAGATAT	ACATATGTCC	CCTATACTAG
121	GTTATTGGAA	AAATTAAGGC	CTTGTGCAAC	CCACTCGACT	TCTTTTGGA	TATCTTGAAG
181	AAAAATATGA	AGAGCATTG	TATGAGCGCG	ATGAAGGTGA	TAAATGGCGA	AACAAAAAGT
241	TTGAATTGGG	TTTGGAGTTT	CCCAATCTTC	CTTATTATAT	TGATGGTGAT	GTTAAATTAA
301	CACAGTCTAT	GGCCATCATA	CGTTATATAG	CTGACAAGCA	CAACATGTTG	GGTGGTTGTC
361	CAAAAGAGCG	TGCAGAGATT	TCAATGCTTG	AAGGAGCGGT	TTTGGATATT	AGATACGGTG
421	TTTCGAGAAT	TGCATATAGT	AAAGACTTTG	AAACTCTCAA	AGTTGATTTT	CTTAGCAAGC
481	TACCTGAAAT	GCTGAAAATG	TTCGAAGATC	GTTTATGTCA	TAAAACATAT	TTAAATGGTG
541	ATCATGTAAC	CCATCCTGAC	TTCATGTTGT	ATGACGCTCT	TGATGTTGTT	TTATACATGG
601	ACCCAATGTG	CCTGGATGCG	TTCCCAAAAT	TAGTTTGT	TAAAAAACGT	ATTGAAGCTA
661	TCCCACAAAT	TGATAAGTAC	TTGAAATCCA	GCAAGTATAT	AGCATGGCCT	TTGCAGGGCT
721	GGCAAGCCAC	GTTTGGTGGT	GGCGACCATC	CTCCAAATC	GGATCTGGTT	CCGCGTCCAT
781	GGTCGAATCA	AACAAGTTTG	TACAAAAAAG	CTGAACGAGA	AACGTAAAAT	GATATAAATA
841	TCAATATATT	AAATTAGATT	TTGCATAAAA	AACAGACTAC	ATAATACTGT	AAAACACAAC
901	ATATCCAGTC	ACTATGGCGG	CCGCATTAGG	CACCCAGGC	TTTACACTTT	ATGCTTCCGG
961	CTCGTATAAT	GTGTGGATTT	TGAGTTAGGA	TCCGTCGAGA	TTTTCAGGAG	CTAAGGAAGC
1021	TAAAATGGAG	AAAAAATCA	CTGGATATAC	CACCGTTGAT	ATATCCCAAT	GGCATCGTAA
1081	AGAACATTTT	GAGGCATTTT	AGTCAGTTGC	TCAATGTACC	TATAACCAGA	CCGTTTCAGCT
1141	GGATATTACG	GCCTTTTAA	AGACCGTAA	GAAAAATAAG	CACAAGTTTT	ATCCGGCCTT
1201	TATTCACATT	CTTGCCCGCC	TGATGAATGC	TCATCCGGAA	TTCCGTATGG	CAATGAAAGA
1261	CGGTGAGCTG	GTGATATGGG	ATAGTGTTCA	CCCTTGTTAC	ACCGTTTCC	ATGAGCAAAC
1321	TGAAACGTTT	TCATCGCTCT	GGAGTGAATA	CCACGACGAT	TTCCGGCAGT	TTCTACACAT
1381	ATATTGCGAA	GATGTGGCGT	GTTACGGTGA	AAACCTGGCC	TATTTCCCTA	AAGGGTTTAT
1441	TGAGAATATG	TTTTTCGTCT	CAGCCAATCC	CTGGGTGAGT	TTCAACAGTT	TTGATTTAAA
1501	CGTGGCCAAT	ATGGACAAT	TCTTCGCCCC	CGTTTTCAAC	ATGGGCAAAT	ATTATACGCA
1561	AGGCGACAAG	GTGCTGATGC	CGCTGGCGAT	TCAGGTTTCA	CATGCCGTCT	GTGATGGCTT
1621	CCATGTCCGC	AGAATGCTTA	ATGAATTACA	ACAGTACTGC	GATGAGTGGC	AGGGCGGGGC
1681	GTAATCTAGA	GGATCCGGCT	TACTAAAAGC	CAGATAACAG	TATGCGTATT	TGCGCGCTGA
1741	TTTTTTCGGT	ATAAGAATAT	ATACTGATAT	GTATACCCGA	AGTATGTCAA	AAAGAGGTGT
1801	GCTATGAAGC	AGCGTATTAC	AGTGACAGTT	GACAGCGACA	GCTATCAGTT	GCTCAAGGCA
1861	TATATGATGT	CAATATCTCC	GGTCTGGTAA	GCACAACCAT	GCAGAATGAA	GCCCCGTCGT
1921	TGCGTGCCGA	ACGCTGGAAA	GCGGAAAATC	AGGAAGGGAT	GGCTGAGGTC	GCCCCGTTTA
1981	TTGAAATGAA	CGGCTCTTTT	GCTGACGAGA	ACAGGGACTG	GTGAAATGCA	GTTTAAGGTT
2041	TACACCTATA	AAAGAGAGAG	CCGTTATCGT	CTGTTTGTGG	ATGTACAGAG	TGATATTATT
2101	GACACGCCCC	GGCGACGGAT	GGTGATCCCC	CTGGCCAGTG	CACGTCTGCT	GTCAGATAAA
2161	GTCTCCCGTG	AACTTTACCC	GGTGGTGCAT	ATCGGGGATG	AAAGCTGGCG	CATGATGACC
2221	ACCGATATGG	CCAGTGTGCC	GGTCTCCGTT	ATCGGGGAAG	AAGTGGCTGA	TCTCAGCCAC
2281	CGCGAAAATG	ACATCAAAAA	CGCCATTAA	CTGATGTTCT	GGGGAATATA	AATGTCAGGC
2341	TCCCTTATAC	ACAGCCAGTC	TGCAGGTCGA	CCATAGTGAC	TGGATATGTT	GTGTTTTACA
2401	GTATTATGTA	GTCTGTTTTT	TATGCAAAAT	CTAATTTAAT	ATATTGATAT	TTATACCAAT
2461	TTACGTTTCT	CGTTCAGCTT	TCTTGACAA	AGTGGTTTGA	TTCGACCCGG	GATCCGGCTG
2521	CTAACAAAGC	CCGAAAGGAA	GCTGAGTTGG	CTGCTGCCAC	CGCTGAGCAA	TAAGTAGCAT
2581	AACCCCTTGG	GGCCTCTAAA	CGGGTCTTGA	GGGGTTTTTT	GCTGAAAGGA	GGAAGTATAT
2641	CCGGATATCC	ACAGGACGGG	TGTGGTCGCC	ATGATCGCGT	AGTCGATAGT	GGCTCCAAGT-

FIGURE 35B

89/240

2701 AGCGAAGCGA GCAGGACTGG GCGGCGGCCA AAGCGGTCGG ACAGTGCTCC GAGAACGGGT
2761 GCGCATAGAA ATTGCATCAA CGCATATAGC GCTAGCAGCA CGCCATAGTG ACTGGCGATG
2821 CTGTCGGAAT GGACGATATC CCGCAAGAGG CCCGCGAGTA CCGGCATAAC CAAGCCTATG
2881 CCTACAGCAT CCAGGGTGAC GGTGCCGAGG ATGACGATGA GCGCATTGTT AGATTTTCATA
2941 CACGGTGCCT GACTGCGTTA GCAATTTAAC TGTGATAAAC TACCGCATT AAGCTTATCG
3001 ATGATAAGCT GTCAAACATG AGAATCTTTG AAGACGAAAAG GGCCTCGTGA TACGCCTATT
3061 TTTATAGGTT AATGTCATGA TAATAATGGT TTCTTAGACG TCAGGTGGCA CTTTTCGGGG
3121 AAATGTGCGC GGAACCCCTA TTTGTTTATT TTTCTAAATA CATTCAAATA TGTATCCGCT
3181 CATGAGACAA TAACCCTGAT AAATGCTTCA ATAATATTGA AAAAGGAAGA GTATGAGTAT
3241 TCAACATTTT CGTGTGCGCC TTATTCCCTT TTTTGC GGCA TTTTGCCTTC CTGTTTTTGC
3301 TCACCCAGAA ACGCTGGTGA AAGTAAAGA TGCTGAAGAT CAGTTGGGTG CACGAGTGGG
3361 TTACATCGAA CTGGATCTCA ACAGCGGTAA GATCCTTGAG AGTTTTCGCC CCGAAGAACG
3421 TTTTCCAATG ATGAGCACTT TTAAAGTTCT GTCGCGGCT ACACTATTCT CAGAATGACT TGGTTGAGTA
3481 CGCCGGGCAA GAGCAACTCG GTCGCGGCT ACACTATTCT CAGAATGACT TGGTTGAGTA
3541 CTCACAGTC ACAGAAAAGC ATCTTACGGA TGGCATGACA GTAAGAGAAT TATGCAGTGC
3601 TGCCATAACC ATGAGTGATA ACACTGCGGC CAACTTACTT CTGACAACGA TCGGAGGACC
3661 GAAGGAGCTA ACCGCTTTTT TGCACAACAT GGGGGATCAT GTAACCTGCC TTGATCGTTG
3721 GGAACCGGAG CTGAATGAAG CCATACCAA CGACGAGCGT GACACCACGA TGCCTGCAGC
3781 AATGGCAACA ACGTTGCGCA AACTATTAAC TGGCGAACTA CTTACTCTAG CTTCCCGGCA
3841 ACAATTAATA GACTGGATGG AGGCGGATAA AGTTGCAGGA CCACTTCTGC GCTCGGCCCT
3901 TCCGGCTGGC TGGTTTATTG CTGATAAATC TGGAGCCGGT GAGCGTGGGT CTCGCGGTAT
3961 CATTGCAGCA CTGGGGCCAG ATGGTAAGCC CTCCCGTATC GTAGTTATCT ACACGACGGG
4021 GAGTCAGGCA ACTATGGATG AACGAAATAG ACAGATCGCT GAGATAGGTG CCTCACTGAT
4081 TAAGCATTGG TAACTGTCAG ACCAAGTTTA CTCATATATA CTTTAGATTG ATTTAAAACT
4141 TCATTTTAA TTTAAAAGGA TCTAGGTGAA GATCCTTTTT GATAATCTCA TGACCAAAAT
4201 CCCTTAACGT GAGTTTTCTG TCCACTGAGC GTCAGACCCC GTAGAAAAGA TCAAAGGATC
4261 TTCTTGAGAT CCTTTTTTTC TGCGCGTAAT CTGCTGCTTG CAAACAAAAA AACCACCGCT
4321 ACCAGCGGTG GTTTGTTTGC CGGATCAAGA GCTACCAACT CTTTTTCCGA AGGTAACCTGG
4381 CTTACGAGCA GCGCAGATAC CAAATACTGT CCTTCTAGTG TAGCCGTAGT TAGGCCACCA
4441 CTTCAAGAAC TCTGTAGCAC CGCCTACATA CCTCGCTCTG CTAATCCTGT TACCAGTGGC
4501 TGCTGCCAGT GCGGATAAGT CGTGTCTTAC CGGTTGGAC TCAAGACGAT AGTTACCGGA
4561 TAAGGCGCAG CGGTGCGGCT GAACGGGGGG TTCGTGCACA CAGCCAGCT TGGAGCGAAC
4621 GACCTACACC GAACGAGAT ACCTACAGCG TGAGCTATGA GAAAGCGCCA CGCTTCCCGA
4681 AGGGAGAAAG GCGGACAGGT ATCCGGTAAG CGGCAGGGTC GGAACAGGAG AGCGCACGAG
4741 GGAGCTTCCA GGGGAAAACG CCTGGTATCT TTATAGTCTT GTCGGGTTTC GCCACCTCTG
4801 ACTTGAGCGT CGATTTTGTG GATGCTCGTC AGGGGGGGCG AGCCTATGGA AAAACGCCAG
4861 CAACGCGGCC TTTTACGGT TCCTGGCCTT TTGCTGGCCT TTTGCTCACA TGTCTTTTCC
4921 TGCGTTATCC CCTGATTCTG TGGATAACCG TATTACCGCC TTTGAGTGAG CTGATACCGC
4981 TCGCCGCAGC CGAACGACCG AGCGCAGCGA GTCAGTGAGC GAGGAAGCGG AAGAGCGCCT
5041 GATGCGGTAT TTTCTCCTTA CGCATCTGTG CGGTATTTCA CACCGCATAT ATGGTGCATC
5101 CTCAGTACAA TCTGCTCTGA TGCCGCATAG TTAAGCCAGT ATACACTCCG CTATCGCTAC
5161 GTGACTGGGT CATGGCTGCG CCCCACACC CGCCAACACC CGCTGACCGG CCCTGACGGG
5221 CTTGCTGCTT CCGGCATCC GCTTACAGAC AAGCTGTGAC CGTCTCCGGG AGCTGCATGT
5281 GTCAGAGGTT TTCACCGTCA TCACCGAAAC GCGCGAGGCA GCTGCGGTAA AGCTCATCAG
5341 CGTGGTCTGT AAGCGATTCA CAGATGTCTG CCTGTTTCATC CGCGTCCAGC TCGTTGAGTT
5401 TCTCCAGAAG CGTTAATGTC TGGCTTCTGA TAAAGCGGGC CATGTTAAGG GCGGTTTTTT
5461 CCTGTTTTGGT CACTGATGCC TCCGTGTAAG GGGGATTTCT GTTCATGGGG GTAATGATAC
5521 CGATGAAACG AGAGAGGATG CTCACGATAC GGGTACTGA TGATGAACAT GCCCGGTTAC
5581 TGGAACGTTG TGAGGGTAAA CAACTGGCGG TATGGATGCG GCGGGACCAG AGAAAAATCA
5641 CTCAGGGTCA ATGCCAGCGC TTCGTTAATA CAGATGTAGG TGTTCCACAG GGTAGCCAGC
5701 AGCATCTGTC GATGCAGATC CGGAACATAA TGGTGACGGG CGCTGACTTC CGCGTTTCCA
5761 GACTTTACGA AACACGGAAA CCGAAGACCA TTCATGTTGT TGCTCAGGTC GCAGACGTTT
5821 TGCAGCAGCA GTCGCTTCAC GTTCGCTCGC GTATCGGTGA TTCATTCTGC TAACCAGTAA
5881 GGCAACCCCG CCAGCCTAGC CGGGTCCTCA ACGACAGGAG CACGATCATG CGCACCCTGT
5941 GCCAGGACCC AACGCTGCCC GAGATGCGCC GCGTGCGGCT GCTGGAGATG GCGGACGCGA
6001 TGGATATGTT CTGCCAAGGG TTGGTTTTCG CATTACAGT TCTCCGCAAG AATTGATTGG
6061 CTCCAATTCT TGGAGTGGTG AATCCGTTAG CGAGGTGCCG CCGGCTTCCA TTCAGSTCGA
6121 GGTGGCCCGG CTCCATGCAC CGCGACGCAA CGCGGGGAGG CAGACAAGGT ATAGGGCGGC-

FIGURE 35C

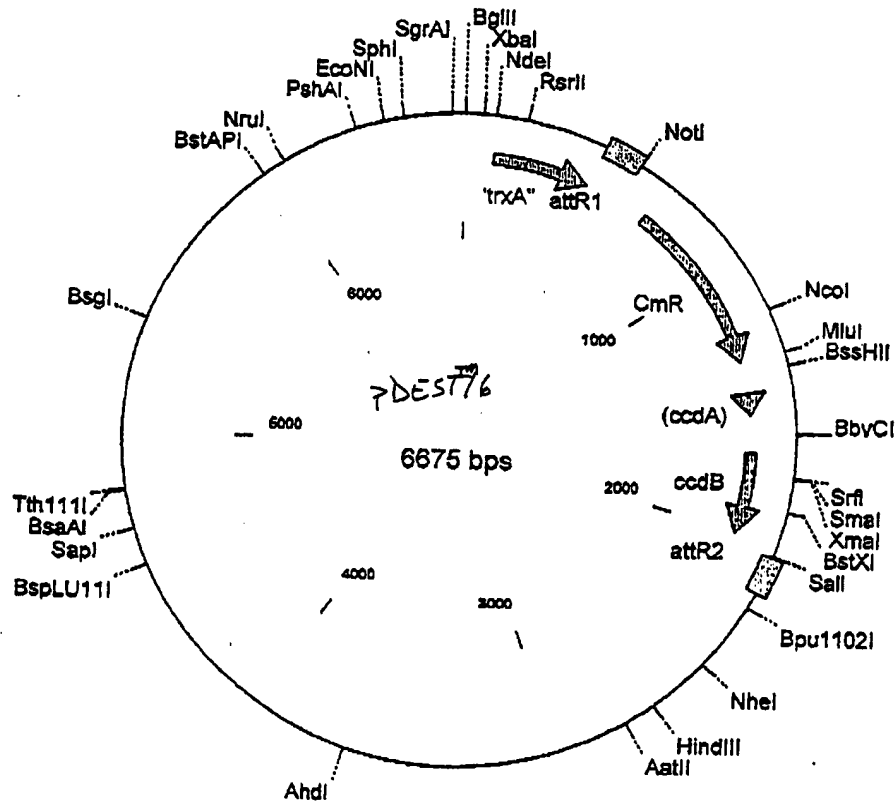
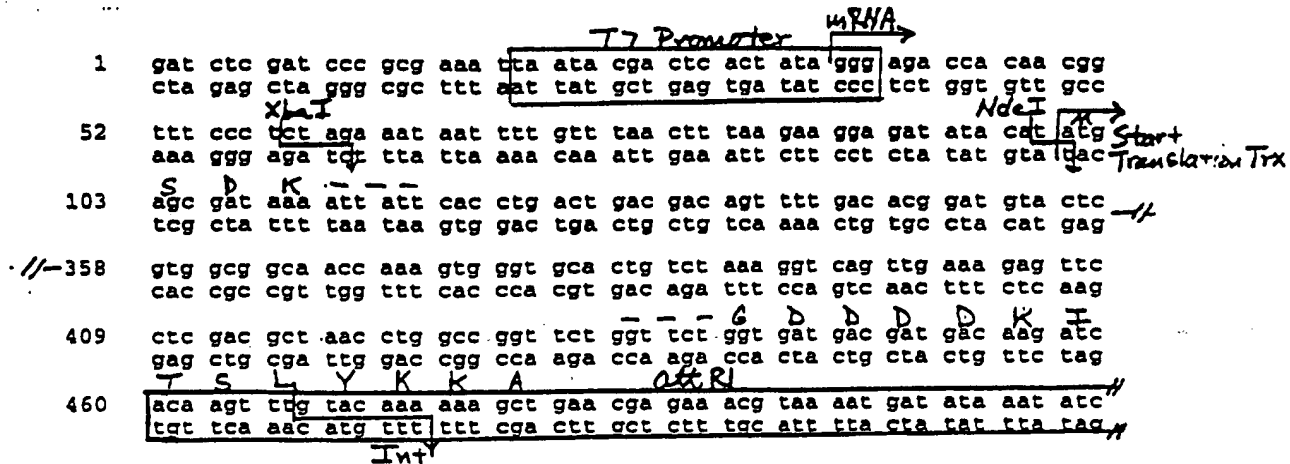
90/240

```
6181 GCCTACAATC CATGCCAACC CGTTCCATGT GCTCGCCGAG GCGGCATAAA TCGCCGTGAC
6241 GATCAGCGGT CCAGTGATCG AAGTTAGGCT GGTAAGAGCC GCGAGCGATC CTTGAAGCTG
6301 TCCCTGATGG TCGTCATCTA CCTGCCTGGA CAGCATGGCC TGCAACGCGG GCATCCCGAT
6361 GCCGCCGGAA GCGAGAAGAA TCATAATGGG GAAGGCCATC CAGCCTCGCG TCGCGAACGC
6421 CAGCAAGACG TAGCCCAGCG CGTCGGCCGC CATGCCGGCG ATAATGGCCT GCTTCTCGCC
6481 GAAACGTTTG GTGGCGGGAC CAGTGACGAA GGCTTGAGCG AGGGCGTGCA AGATTCCGAA
6541 TACCGCAAGC GACAGGCCGA TCATCGTCGC GCTCCAGCGA AAGCGGTCCT CGCCGAAAAA
6601 GACCCAGAGC GCTGCCGGCA CCTGTCCTAC GAGTTGCATG ATAAAGAAGA CAGTCATAAG
6661 TGCGGCGACG ATAGTCATGC CCCGCGCCCA CCGGAAGGAG CTGACTGGGT TGAAGGCTCT
6721 CAAGGGCATC GGTGATCGA CGCTCTCCCT TATGCGACTC CTGCATTAGG AAGCAGCCCA
6781 GTAGTAGGTT GAGGCCGTTG AGCACCGCCG CCGCAAGGAA TGGTGCATGC AAGGAGATGG
6841 CGCCCAACAG TCCCCCGGCC ACGGGGCCTG CCACCATACC CACGCCGAAA CAAGCGCTCA
6901 TGAGCCCGAA GTGGCGAGCC CGATCTTCCC CATCGGTGAT GTCGGCGATA TAGGCGCCAG
6961 CAACCGCACC TGTGGCGCCG GTGATGCCGG CCACGATGCG TCCGGCGTAG AGG
```

FIGURE 351)

91/240

Figure 36A: γ DEST16 **Thioredoxin N-Fusion Protein**
in *E. coli* with T7 Promoter



92/240

pDEST16 6675 bp

<u>Location (Base Nos.)</u>			<u>Gene Encoded</u>			
104..457			trxA			
585..461			attR1			
694..1353			CmR			
1473..1557			inactivated ccdA			
1695..2000			ccdB			
2041..2165			attR2			
1	AGATCTCGAT	CCCGCGAAAT	TAATACGACT	CACTATAGGG	AGACCACAAC	GGTTTCCCTC
61	TAGAAATAAT	TTTGTTTAAC	TTTAAGAAAG	AGATATACAT	ATGAGCGATA	AAATTATTCA
121	CCTGACTGAC	GACAGTTTTG	ACACGGATGT	ACTCAAAGCG	GACGGGGCGA	TCTTCGTCGA
181	TTTCTGGGCA	GAGTGGTGCG	GTCCGTGCAA	AATGATCGCC	CCGATTCTGG	ATGAAATCGC
241	TGACGAATAT	CAGGGCAAAC	TGACCGTTGC	AAAACCTGAAC	ATCGATCAAA	ACCCTGGCAC
301	TGCGCCGAAA	TATGGCATCC	GTGGTATCCC	GACTCTGCTG	CTGTTCAAAA	ACGGTGAAGT
361	GGCGGCAACC	AAAGTGGGTG	CACTGTCTAA	AGGTCAGTTG	AAAGAGTTCC	TCGACGCTAA
421	CCTGGCCGGT	TCTGGTTCTG	GTGATGACGA	TGACAAGATC	ACAAGTTTGT	ACAAAAAAGC
481	TGAACGAGAA	ACGTAAAATG	ATATAAATAT	CAATATATTA	AATTAGATTT	TGCATAAAAA
541	ACAGACTACA	TAATACTGTA	AAACACAACA	TATCCAGTCA	CTATGGCGGC	CGCATTAGGC
601	ACCCGAGGCT	TTACACTTTA	TGCTTCCGGC	TCGTATAATG	TGTGGATTTT	GAGTTAGGAT
661	CCGGCGAGAT	TTTCAGGAGC	TAAGGAAGCT	AAAATGGAGA	AAAAAATCAC	TGGATATACC
721	ACCGTTGATA	TATCCCAATG	GCATCGTAAA	GAACATTTTG	AGGCATTTCA	GTCACTTGCT
781	CAATGTACCT	ATAACCAGAC	CGTTCAGCTG	GATATTACGG	CCTTTTTTAA	GACCGTAAAG
841	AAAAATAAGC	ACAAGTTTTA	TCCGGCCTTT	ATTACATTTC	TTGCCCGCCT	GATGAATGCT
901	CATCCGGAAT	TCCGTATGGC	AATGAAAGAC	GGTGAGCTGG	TGATATGGGA	TAGTGTTCAC
961	CCTTGTTACA	CCGTTTTCCA	TGAGCAAAC	GAAACGTTTT	CATCGCTCTG	GAGTGAATAC
1021	CACGACGATT	TCCGGCAGTT	TCTACACATA	TATTGCGAAG	ATGTGGCGTG	TTACGGTGAA
1081	AACCTGGCCT	ATTTCCCTAA	AGGGTTTATT	GAGAATATGT	TTTTCGTCTC	AGCCAATCCC
1141	TGGGTGAGTT	TCACCAAGTT	TGATTTAAAC	GTGGCCAATA	TGGACAACCT	CTTCGCCCCC
1201	GTTTTACCA	TGGGCAAATA	TTATACGCAA	GGCGACAAGG	TGCTGATGCC	GCTGGCGATT
1261	CAGGTTTCATC	ATGCCGTCTG	TGATGGCTTC	CATGTCGGCA	GAATGCTTAA	TGAATTACAA
1321	CAGTACTGCG	ATGAGTGGCA	GGCGGGGGCG	TAAACGCGTG	GATCCGGCTT	ACTAAAAGCC
1381	AGATAACAGT	ATGCGTATTT	GCGCGCTGAT	TTTTGCGGTA	TAAGAATATA	TACTGATATG
1441	TATACCCGAA	GTATGTCAAA	AAGAGGTGTG	CTATGAAGCA	GCGTATTACA	GTGACAGTTG
1501	ACAGCGACAG	CTATCAGTTG	CTCAAGGCAT	ATATGATGTC	AATATCTCCG	GTCTGGTAAG
1561	CACAACCATG	CAGAATGAAG	CCCGTCGTCT	CGGTGCCGAA	CGCTGGAAAG	CGGAAAATCA
1621	GGAGAGGATG	GCTGAGGTCG	CCCGGTTTAT	TGAAATGAAC	GGCTCTTTTG	CTGACGAGAA
1681	CAGGGACTGG	TGAAATGCAG	TTTAAGGTTT	ACACCTATAA	AAGAGAGAGC	CGTTATCGTC
1741	TGTTTGTGGA	TGTACAGAGT	GATATTATTG	ACACGCCCCG	GCGACGGATG	GTGATCCCCC
1801	TGGCCAGTGC	ACGTCTGCTG	TCAGATAAAG	TCTCCCGTGA	ACTTTACCCG	GTGGTGCATA
1861	TCGGGGATGA	AAGCTGGCGC	ATGATGACCA	CCGATATGGC	CAGTGTGCCG	GTCTCCGTTA
1921	TCGGGGAAGA	AGTGGCTGAT	CTCAGCCACC	GCGAAAATGA	CATCAAAAAC	GCCATTAACC
1981	TGATGTTCTG	GGGAATATAA	ATGTCAGGCT	CCCTTATACA	CAGCCAGTCT	GCAGGTCGAC
2041	CATAGTGAAT	GGATATGTTG	TGTTTTACAG	TATTATGTAG	TCTGTTTTTT	ATGCAAAATC
2101	TAATTTAATA	TATTGATATT	TATATCATT	TACGTTTCTC	GTTTACGCTT	CTTGATACAAA
2161	GTGGTGATGA	TCCGGCTGCT	AACAAAGCCC	GAAAGGAAGC	TGAGTTGGCT	GCTGCCACCG
2221	CTGAGCAATA	ACTAGCATAA	CCCCTGGGGG	CCTCTAAACG	GGTCTTGAGG	GGTTTTTTTG
2281	TGAAAGGAGG	AACTATATCC	GGATATCCAC	AGGACGGGTG	TGGTCGCCAT	GATCGCGTAG
2341	TCGATAGTGG	CTCCAAGTAG	CGAAGCGAGC	AGGACTGGGC	GGCGGCCAAA	GCGGTCCGGAC
2401	AGTGCTCCGA	GAACGGGTGC	GCATAGAAAT	TGCATCAACG	CATATAGCGC	TAGCAGCACG
2461	CCATAGTGAC	TGGCGATGCT	GTCCGAATGG	ACGATATCCC	GCAAGAGGCC	CGGCAGTACC
2521	GGCATAACCA	AGCCTATGCC	TACAGCATCC	AGGGTGACGG	TGCCGAGGAT	GACGATGAGC
2581	GCATTGTTAG	ATTTTCATACA	CGGTGCCTGA	CTGCGTTAGC	AATTTAACTG	TGATAAACTA
2641	CCGCATTAAA	GCTTATCGAT	GATAAGCTGT	CAAACATGAG	AATTCTTGAA	GACGAAAGGG
2701	CCTCGTGATA	CGCCTATTTT	TATAGGTTAA	TGTCATGATA	ATAATGGTTT	CTTAGACGTC
2761	AGGTGGCACT	TTTCGGGGAA	ATGTGCGCGG	AACCCCTATT	TGTTTATTTT	TCTAAATACA-

FIGURE 36B

93/240

2821 TTCAAATATG TATCCGCTCA TGAGACAATA ACCCTGATAA ATGCTTCAAT AATATTGAAA
 2881 AAGGAAGAGT ATGAGTATTC AACATTTCCG TGTCGCCCTT ATTCCCTTTT TTGCGGCATT
 2941 TTGCCTTCCT GTTTTGTCTC ACCCAGAAAC GCTGGTGAAA GTAAAAGATG CTGAAGATCA
 3001 GTTGGGTGCA CGAGTGGGTT ACATCGAACT GGATCTCAAC AGCGGTAAAG TCCTTGAGAG
 3061 TTTTCGCCCC GAAGAACGTT TTCCAATGAT GAGCACTTTT AAAGTTCTGC TATGTGGCGC
 3121 GGTATTATCC CGTGTGACG CCGGGCAAGA GCAACTCGGT CGCCGCATAC ACTATTCTCA
 3181 GAATGACTTG GTTGAGTACT CACCACTCAC AGAAAAGCAT CTTACGGATG GCATGACAGT
 3241 AAGAGAATTA TGCAGTGCTG CCATAACCAT GAGTGATAAC ACTGCGGCCA ACTTACTTCT
 3301 GACAACGATC GGAGGACCGA AGGAGCTAAC CGCTTTTTTG CACAACATGG GGGATCATGT
 3361 AACTCGCCTT GATCGTTGGG AACCGGAGCT GAATGAAGCC ATACCAAACG ACGAGCGTGA
 3421 CACCACGATG CCTGCAGCAA TGGCAACAAC GTTGCGCAAA CTATTAAGTG CGGAAGTACT
 3481 TACTCTAGCT TCCCGGCAAC AATTAATAGA CTGGATGGAG GCGGATAAAG TTGCAGGACC
 3541 ACTTCTGCGC TCGGCCCTTC CGGCTGGCTG GTTTATTGCT GATAAATCTG GAGCCGGTGA
 3601 CGGTGGGTCT CGCGGTATCA TTGCAGCACT GGGGCCAGAT GGTAAGCCCT CCCGTATCGT
 3661 AGTTATCTAC ACGACGGGGA GTCAGGCAAC TATGGATGAA CGAAATAGAC AGATCGCTGA
 3721 GATAGGTGCC TCACTGATTA AGCATTGGTA ACTGTCAGAC CAAGTTTACT CATATATACT
 3781 TTAGATTGAT TTAAAACTTC ATTTTAAAT TAAAAGGATC TAGGTGAAGA TCCTTTTGA
 3841 TAATCTCATG ACCAAAATCC CTTAACGTGA GTTTTCGTTC CACTGAGCGT CAGACCCCGT
 3901 AGAAAAGATC AAAGGATCTT CTTGAGATCC TTTTTTCTG CGCGTAATCT GCTGCTTGCA
 3961 AACAAAAAAA CCACCGCTAC CAGCGGTGGT TTGTTTGCCG GATCAAGAGC TACCAACTCT
 4021 TTTTCCGAAG GTAACCTGGT TCAGCAGAGC GCAGATACCA AATACTGTCC TTCTAGTGA
 4081 GCCGTAGTTA GGCCACCACT TCAAGAAGCT TGTAGCACCG CCTACATACC TCGCTCTGCT
 4141 AATCCTGTTA CCAGTGGCTG CTGCCAGTGG CGATAAGTCG TGTCTTACCG GGTGGAAGTC
 4201 AAGACGATAG TTACCGGATA AGGCGCAGCG GTCGGGCTGA ACGGGGGGTT CGTGCACACA
 4261 GCCCAGCTTG GAGCGAACGA CCTACACCGA ACTGAGATAC CTACAGCGTG AGCTATGAGA
 4321 AAGCGCCACG CTTCCCGAAG GGAGAAAGGC GGACAGGTAT CCGGTAAGCG GCAGGGTCGG
 4381 AACAGGAGAG CGCACGAGGG AGCTTCCAGG GGGAAACGCC TGGTATCTTT ATAGTCCTGT
 4441 CGGGTTTCGC CACCTCTGAC TTGAGCGTCG ATTTTGTGA TGCTCGTCAG GGGGGCGGAG
 4501 TCCGGTGAAG AACGCCAGCA ACGCGGCCCT TTTACGGTTC CTGGCCTTTT GCTGGCCTTT
 4561 TGCTCACATG TTCTTCTCTG CGTTATCCCC TGATTCTGTG GATAACCGTA TTACCGCCTT
 4621 TGAGTGAGCT GATACCGCTC GCCGCAGCCG AACGACCGAG CGCAGCGAGT CAGTGAGCGA
 4681 GGAAGCGGAA GAGCGCCTGA TGCGGTATTT TCTCCTTACG CATCTGTGCG GTATTTTACA
 4741 CCGCATATAT GGTGCACTCT CAGTACAATC TGCTCTGATG CCGCATAGTT AAGCCAGTAT
 4801 ACACTCCGCT ATCGCTACGT GACTGGGTCA TGGCTGCGCC CCGACACCCG CCAACACCCG
 4861 CTGACGCGCC CTGACGGGCT TGTCTGCTCC CGGCATCCGC TTACAGACAA GCTGTGACCG
 4921 TCTCCGGGAG CTGCATGTGT CAGAGGTTTT CACCGTCATC ACCGAAACGC GCGAGGCGAG
 4981 TCCGGTAAAG CTCATCAGCG TGGTCTGTAA GCGATTACCA GATGTCTGCC GTTTCATCCG
 5041 CGTCCAGCTC GTTGAGTTTC TCCAGAAGCG TTAATGTCTG GCTTCTGATA AAGCGGGCCA
 5101 TGTTAAGGGC GGTTTTTTCC TGTTTGGTCA CTGATGCCTC CGTGTAAAGG GGATTTCTGT
 5161 TCATGGGGGT AATGATACCG ATGAAACGAG AGAGGATGCT CACGATACGG GTTACTGATG
 5221 ATGAACATGC CCGGTTACTG GAACGTTGTG AGGGTAAACA ACTGGCGGTA TGGATGCGGC
 5281 GGGACCAGAG AAAAATCACT CAGGGTCAAT GCCAGCGCTT CGTTAATACA GATGTAGGTG
 5341 TTCCACAGGG TAGCCAGCAG CATCCTGCGA TGCAGATCCG GAACATAATG GTGCAGGCG
 5401 CTGACTTCCG CGTTTCCAGA CTTTACGAAA CACGGAAAAC GAAGACCATT CATGTTGTTG
 5461 CTCAGGTGCG AGACGTTTTG CAGCAGCAGT CGCTTCACGT TCGCTCGCGT ATCGGTGATT
 5521 CATTCTGCTA ACCAGTAAGG CAACCCCGCC AGCCTAGCCG GGTCTCAAC GACAGGAGCA
 5581 CGATCATGCG CACCCGTGGC CAGGACCCAA CGCTGCCCGA GATGCGCCGC GTGCGGCTGC
 5641 TGGAGATGGC GGACCGCATG GATATGTTCT GCCAAGGGTT GGTTCGCGCA TTCACAGTTC
 5701 TCCGCAAGAA TTGATTGGCT CCAATTCTTG GAGTGGTGAA TCCGTTAGCG AGGTGCCGCC
 5761 GGCTTCCATT CAGGTCGAGG TGGCCCGGCT CCATGCACCG CGACGCAACG CCGGGAGGCA
 5821 GACAAGGTAT AGGGCGGCGC TCACAATCCA TGCCAACCCG TTCCATGTGC TCGCCGAGGC
 5881 GGCATAAATC GCCGTGACGA TCAGCGGTCC AGTGATCGAA GTTAGGCTGG TAAGAGCCGC
 5941 GAGCGATCCT TGAAGCTGTC CCTGATGGTC GTCATCTACC TGCCTGGACA GCATGGCCTG
 6001 CAACGCGGGC ATCCCGATGC CGCCGGAAGC GAGAAGAATC ATAATGGGGA AGGCCATCCA
 6061 GCCTCGCCTC GCGAACGCCA GCAAGACGTA GCCCAGCGCG TCGGCCGCCA TGCCGCGGAT
 6121 AATGGCCTGC TTCTCGCCGA AACGTTTGGT GGCGGGACCA GTGACGAAGG CTTGAGCGAG
 6181 GCGGTGCAAG ATTCGAATA CCGCAAGCGA CAGGCCGATC ATCGTCGCGC TCCAGCGAAA
 6241 GCGTCTCTCG CCGAAAATGA CCCAGAGCGC TGCCGGCACC TGTCCTACGA GTTGATGATG

FIGURE 36C

94/240

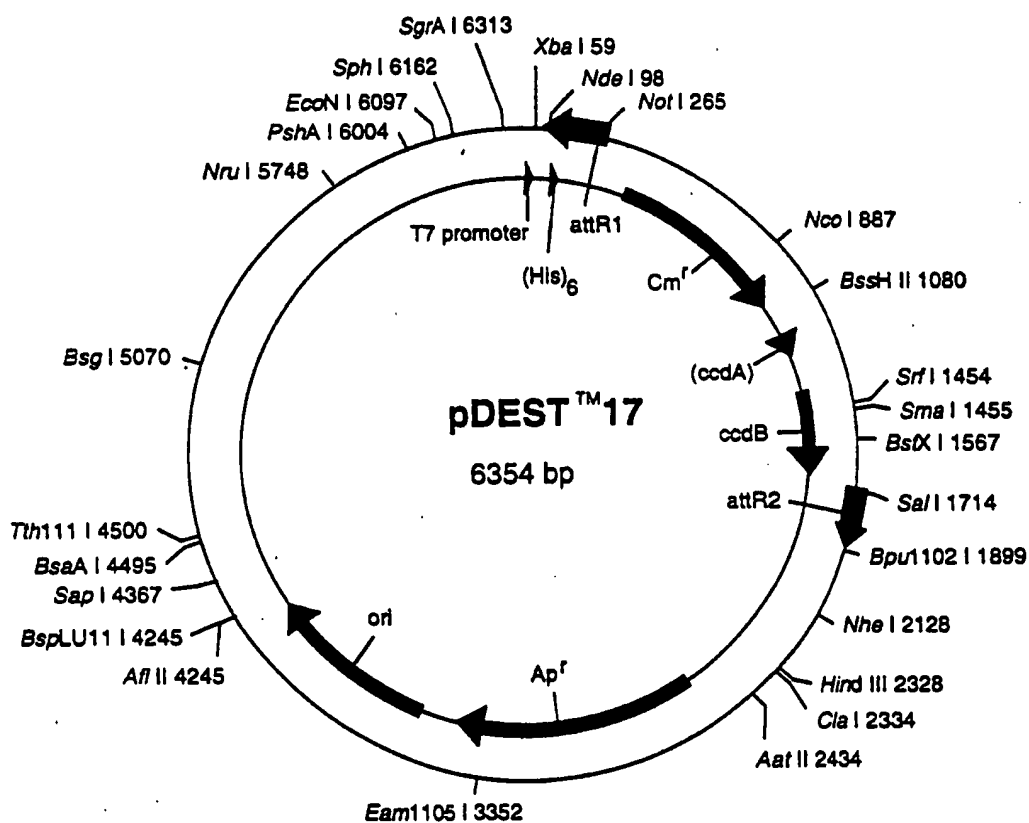
6301 AAAGAAGACA GTCATAAGTG CGGCGACGAT AGTCATGCCC CGCGCCCACC GGAAGGAGCT
6361 GACTGGGTTG AAGGCTCTCA AGGGCATCGG TCGATCGACG CTCTCCCTTA TGCGACTCCT
6421 GCATTAGGAA GCAGCCCAGT AGTAGGTTGA GGCCGTTGAG CACCGCCGCC GCAAGGAATG
6481 GTGCATGCAA GGAGATGGCG CCCAACAGTC CCCCGGCCAC GGGGCTGCTG ACCATACCCA
6541 CGCCGAAACA AGCGCTCATG AGCCCGAAGT GGCAGAGCCG ATCTTCCCA TCGGTGATGT
6601 CGGCGATATA GGCGCCAGCA ACCGCACCTG TGGCGCCGGT GATGCCGGCC ACGATGCGTC
6661 CGGCGTAGAG GATCG

FIGURE 36D

95/240

1 gat ccc gcg aaa tca ata cga cto act ata ggg aga cca caa cgg ttt ccc
 cta ggg cgc ttt aat tat gct gag tga tat ccc tct ggt gtt gcc aaa ggg
 52 tct aga aat aat ttt gtt taa ctt taa gaa gga gat ata cat atg tgg tac
 aga tct tta tta aaa caa att gaa att ctt cct cta tat gta tac agc atg
 103 Y H H H H H H L E S T S L Y K K A
 tac cat cac cat cac cat ctc gaa tca aca agt tgg tac aaa aaa gct
 atg gta gtg gta gtg gta gtg gag ctt agt tgt tca aac atg ttt ttt cga
 attR1 Int

T7 Promoter mRNA
 Start Translation



pDEST17 6354 bp

<u>Location (Base Nos.)</u>		<u>Gene Encoded</u>
258..134		attR1
367..1026		CmR
1146..1230		inactivated ccdA
1368..1673		ccdB
1714..1838		attR2
2564..3421		ampR

1	CGATCCCGCG	AAATTAATAC	GACTCACTAT	AGGGAGACCA	CAACGGTTTC	CCTCTAGAAA
61	TAATTTTGTT	TAACTTTAAG	AAGGAGATAT	ACATATGTCG	TACTACCATC	ACCATCACCA
121	TCACCTCGAA	TCAACAAGTT	TGTACAAAAA	AGCTGAACGA	GAAACGTAAA	ATGATATAAA
181	TATCAATATA	TTAAATTAGA	TTTTGCATAA	AAAACAGACT	ACATAATACT	GTA AACACA
241	ACATATCCAG	TCATATGGC	GGCCGCATTA	GGCACCCAG	GCTTTACACT	TTATGCTTCC
301	GGCTCGTATA	ATGTGTGGAT	TTTGAGTTAG	GATCCGTCGA	GATTTTCAGG	AGCTAAGGAA
361	GCTAAATGG	AGAAAAAAT	CACTGGATAT	ACCACCGTTG	ATATATCCCA	ATGGCATCGT
421	AAAGAACATT	TTGAGGCATT	TCAGTCAGTT	GCTCAATGTA	CCTATAACCA	GACCGTTCAG
481	CTGGATATTA	CGGCCTTTTT	AAAGACCGTA	AAGAAAAATA	AGCACAAAGT	TTATCCGGCC
541	TTTATTCACA	TTCTTGCCCG	CCTGATGAAT	GCTCATCCGG	AATTCCGTAT	GGCAATGAAA
601	GACGGTGAGC	TGGTGATATG	GGATAGTGTT	CACCCTTGTT	ACACCGTTTT	CCATGAGCAA
661	ACTGAAACGT	TTTCATCGCT	CTGGAGTGAA	TACCACGACG	ATTTCCGGCA	GTTTCTACAC
721	ATATATTCGC	AAGATGTGGC	GTGTTACGGT	GAAAACCTGG	CCTATTTCCC	TAAAGGGTTT
781	ATTGAGAATA	TGTTTTTCGT	CTCAGCCAAT	CCCTGGGTGA	GTTTCACCAG	TTTTGATTTA
841	AACGTGGCCA	ATATGGACAA	CTTCTTCGCC	CCCGTTTTCA	CCATGGGCAA	ATATTATACG
901	CAAGGCGACA	AGGTGCTGAT	GCCGCTGGCG	ATTCAGGTTT	ATCATGCCGT	CTGTGATGGC
961	TTCCATGTCT	GCAGATGCT	TAATGAATTA	CAACAGTACT	GCGATGAGTG	GCAGGGCGGG
1021	GCGTAAAGAT	CTGGATCCGG	CTTACTAAAA	GCCAGATAAC	AGTATGCGTA	TTTGCGCGCT
1081	GATTTTTGCG	GTATAAGAAT	ATATACTGAT	ATGTATACCC	GAAGTATGTC	AAAAAGAGGT
1141	GTGCTATGAA	GCAGCGTATT	ACAGTGACAG	TTGACAGCGA	CAGCTATCAG	TTGCTCAAGG
1201	CATATATGAT	GTCAATATCT	CCGGTCTGGT	AAGCACAAAC	ATGCAGAATG	AAGCCCCTCG
1261	TCTGCGTGCC	GAACGCTGGA	AAGCGGAAAA	TCAGGAAGGG	ATGGCTGAGG	TCGCCCGGTT
1321	TATTGAAATG	AACGGCTCTT	TTGCTGACGA	GAACAGGGAC	TGGTGAAATG	CAGTTTAAAG
1381	TTTACACCTA	TAAAAGAGAG	AGCCGTTATC	GTCTGTTTGT	GGATGTACAG	AGTGATATTA
1441	TTGACACGCC	CGGGCGACGG	ATGGTGATCC	CCCTGGCCAG	TGCACGTCTG	CTGTCAGATA
1501	AAGTCTCCCG	TGAACTTTAC	CCGGTGGTGC	ATATCGGGGA	TGAAAGCTGG	CGCATGATGA
1561	CCACCGATAT	GGCCAGTGTG	CCGGTCTCCG	TTATCGGGGA	AGAAGTGGCT	GATCTCAGCC
1621	ACCGCGAAAA	TGACATCAAA	AACGCCATTA	ACCTGATGTT	CTGGGGAATA	TAAATGTCAG
1681	GCTCCCTTAT	ACACAGCCAG	TCTGCAGGTC	GACCATAGTG	ACTGGATATG	TTGTGTTTTA
1741	CAGTATTATG	TAGTCTGTTT	TTTATGCAAA	ATCTAATTTA	ATATATTGAT	ATTTATATCA
1801	TTTACGTTT	CTCGTTACAG	TTTCTTGATC	AAAGTGGTTG	ATTCGAGGCT	GCTAACAAAG
1861	CCCGAAAGGA	AGCTGAGTTG	GCTGCTGCCA	CCGCTGAGCA	ATAACTAGCA	TAACCCCTTG
1921	GGGCCTCTAA	ACGGGTCTTG	AGGGGTTTTT	TGCTGAAAGG	AGGAAGTATA	TCCGGATATC
1981	CACAGGACGG	GTGTGGTCGC	CATGATCGCG	TAGTCGATAG	TGGCTCCAAG	TAGCGAAGCG
2041	AGCAGGACTG	GGCGGCGGCC	AAAGCGGTCTG	GACAGTGCTC	CGAGAACGGG	TGCGCATAGA
2101	AATTGCATCA	ACGCATATAG	CGCTAGCAGC	ACGCCATAGT	GA CTGGCGAT	GCTGTGCGAA
2161	TGGACGATAT	CCCGCAAGAG	GCCCGGCAGT	ACCGGCATAA	CCAAGCCTAT	GCCTACAGCA
2221	TCCAGGGTGA	CGGTGCCGAG	GATGACGATG	AGCGCATTGT	TAGATTTTCAT	ACACGGTGCC
2281	TGACTGCGTT	AGCAATTTAA	CTGTGATAAA	CTACCGCATT	AAAGCTTATC	GATGATAAGC
2341	TGTCAAACAT	GAGAATTCTT	GAAGACGAAA	GGGCCTCGTG	ATACGCCTAT	TTTTATAGGT
2401	TAATGTGATG	ATAATAATGG	TTTCTTAGAC	GTCAGGTGGC	ACTTTTCGGG	GAAATGTGCG
2461	CGGAACCCCT	ATTTGTTTAT	TTTTCTAAAT	ACATTCAAAT	ATGTATCCGC	TCATGAGACA
2521	ATAACCCCTGA	TAAATGCTTC	AATAATATTG	AAAAAGGAAG	AGTATGAGTA	TTCAACATTT
2581	CCGTGTCGCC	CTTATTCCCT	TTTTTGCGGC	ATTTTGCCCT	CCTGTTTTTG	CTCACCAGCA
2641	AACGCTGGTG	AAAGTAAAAG	ATGCTGAAGA	TCAGTTGGGT	GCACGAGTGG	GTTACATCGA-

FIGURE 37B

97/240

2701 ACTGGATCTC AACAGCGGTA AGATCCTTGA GAGTTTTCGC CCCGAAGAAC GTTTTCCAAT
2761 GATGAGCACT TTAAAGTTC TGCTATGTGG CGCGGTATTA TCCCGTGTG ACGCCGGGCA
2821 AGAGCAACTC GGTGCGCGCA TACACTATTC TCAGAATGAC TTGGTTGAGT ACTCACCAGT
2881 CACAGAAAAG CATCTTACGG ATGGCATGAC AGTAAGAGAA TTATGCAGTG CTGCCATAAC
2941 CATGAGTGAT AACACTGCGG CCAACTTACT TCTGACAACG ATCGGAGGAC CGAAGGAGCT
3001 AACCGCTTTT TTGCACAACA TGGGGGATCA TGTAACCTCG CTTGATCGTT GGGAAACCGGA
3061 GCTGAATGAA GCCATACCAA ACGACGAGCG TGACACCACG ATGCCTGCAG CAATGGCAAC
3121 AACGTTGCGC AAACATTATA CTGGCGAACT ACTTACTCTA GCTTCCCGGC AACAAATTAAT
3181 AGACTGGATG GAGGCGGATA AAGTTGCAGG ACCACTTCTG CGCTCGGCCC TTCCGGCTGG
3241 CTGGTTTATT GCTGATAAAT CTGGAGCCGG TGAGCGTGGG TCTCGCGGTA TCATTGCAGC
3301 ACTGGGGCCA GATGGTAAGC CCTCCCGTAT CGTAGTTATC TACACGACGG GGAGTCAGGC
3361 AACTATGGAT GAACGAAATA GACAGATCGC TGAGATAGGT GCCTCACTGA TTAAGCATTG
3421 GTAACGTGCA GACCAAGTTT ACTCATATAT ACTTTAGATT GATTTAAAC TTCATTTTAA
3481 ATTTAAAGG ATCTAGGTGA AGATCCTTTT TGATAATCTC ATGACCAAAA TCCCTTAAAG
3541 TGAGTTTTCG TTCCACTGAG CGTCAGACCC CGTAGAAAAG ATCAAAGGAT CTTCTTGAGA
3601 TCCTTTTTTT CTGCGCGTAA TCTGCTGCTT GCAAACAAAA AAACCACCGC TACCAGCGGT
3661 GGTTTGTTTG CCGGATCAAG AGCTACCAAC TCTTTTTCCG AAGGTAACGT GCTTCAGCAG
3721 AGCGCAGATA CCAAATACTG TCCTTCTAGT GTAGCCGTAG TTAGGCCACC ACTTCAAGAA
3781 CTCTGTAGCA CCGCTACAT ACCTCGCTCT GCTAATCCTG TTACCAGTGG CTGCTGCCAG
3841 TGGCGATAAG TCGTGTCTTA CCGGGTTGGA CTCAGACGA TAGTTACCGG ATAAGGCGCA
3901 GCGGTGCGGC TGAACGGGGG GTTCGTGCAC ACAGCCCAGC TTGGAGCGAA CGACCTACAC
3961 CGAACTGAGA TACCTACAGC GTGAGCTATG AGAAAGCGCC ACGCTTCCCG AAGGGAGAAA
4021 GGCGGACAGG TATCCGGTAA GCGGCAGGGT CGGAACAGGA GAGCGCACGA GGGAGCTTCC
4081 AGGGGGAAC GCCTGGTATC TTTATAGTCC TGTGCGGTTT CGCCACCTCT GACTTGAGCG
4141 TCGATTTTTG TGATGCTCGT CAGGGGGGCG GAGCCTATGG AAAACGCCA GCAACGCGGC
4201 CTTTTTACGG TTCTGGCCT TTTGCTGGCC TTTTGCTCAC ATGTTCTTTC CTGCGTTATC
4261 CCCTGATTCT GTGGATAACC GTATTACCGC CTTTGAGTGA GCTGATACCG CTCGCCGAG
4321 CCGAACGACC GAGCGCAGCG AGTCAGTGAG CGAGGAAGCG GAAGAGCGCC TGATGCGGTA
4381 TTTTCTCCTT ACGCATCTGT GCGGTATTTT ACACCGCATA TATGGTGCAC TCTCAGTACA
4441 ATTCGTCTTG ATGCCGCATA GTTAAGCCAG TATACACTCC GCTATCGCTA CGTGAAGTGG
4501 TCATGGCTGC GCGCCGACAC CCGCCAACAC CCGCTGACGC GCGCTGACGG GCTTGTCTGC
4561 TCCCGGCATC CGCTTACAGA CAAGCTGTGA CCGTCTCCGG GAGCTGCATG TGTCAGAGGT
4621 TTTACCGTC ATCACCAGAA CGCGCAGGC AGCTGCGGTA AAGCTCATCA GCGTGGTCGT
4681 GAAGCCATTG ACAGATGTCT GCCTGTTTCT CCGCGTCCAG CTCGTTGAGT TTCTCCAGAA
4741 GCGTTAATGT CTGGCTTCTG ATAAAGCGGG CCATGTTAAG GCGCGTTTTT TCCTGTTTGG
4801 TCACTGATGC CTCCGTGTAA GGGGGATTTC TGTTTCATGG GGTAAATGATA CCGATGAAC
4861 GAGAGAGGAT GCTCAGGATA CGGGTTACTG ATGATGAACA TGCCCGGTTA TGCGAACGTT
4921 GTGAGGGTAA ACAACTGGCG GTATGGATGC GCGGGGACCA GAGAAAAATC ACTCAGGGTC
4981 AATGCCAGCG CTTGCTTAAT ACAGATGTAG GTGTTCCACA GGGTAGCCAG CAGCATCCTG
5041 CGATGCAGAT CCGGAACATA ATGGTGCAGG GCGCTGACTT CCGCGTTTCC AGACTTTACG
5101 AAACACGGAA ACCGAAGACC ATTCATGTTG TTGCTCAGGT CGCAGACGTT TTGCAGCAGC
5161 AGTCGCTTCA CGTTCGCTCG CGTATCGGTG ATTCATTCTG CTAACCAGTA AGGCAACCCC
5221 GCCAGCCTAG CCGGGTCTCT AACGACAGGA GCACGATCAT GCGCACCCGT GGCCAGGACC
5281 CAACGCTGCC CGAGATGCGC CGCGTGCGGC TGCTGGAGAT GGCGGACGCG ATGGATATGT
5341 TCTGCCAAGG GTTGTTTTCG GCATTACAG TTCTCCGCAA GAATTGATTG GCTCCAATTG
5401 TTGGAGTGGT GAATCCGTTA GCGAGGTGCC GCGGCTTCC ATTCAGGTGCG AGGTGGCCCCG
5461 GCTCCATGCA CCGCGACGCA ACGCGGGGAG GCAGACAAGG TATAGGGCGG CGCCTACAAT
5521 CCATGCCAAC CCGTCCATG TGCTCGCCGA GCGGCGATAA ATCGCCGTGA CGATCAGCGG
5581 TCCAGTGATC GAAGTTAGGC TGGTAAGAGC CGCGAGCGAT CCTTGAAGCT GTCCCTGATG
5641 GTCGTCATCT ACCTGCCTGG ACAGCATGGC CTGCAACCGG GGCATCCCGA TGCCGCCGGA
5701 AGCGAGAAGA ATCATAATGG GGAAGGCCAT CCAGCCTCGC GTCGCGAAGC CCAGAACAGC
5761 GTAGCCAGC GCGTCGGCGG CCATGCGCGG GATAATGGCC TGCTTCTCGC CGAAACGTTT
5821 GGTGGCGGGA CCACTGACGA AGGCTTGAGC GAGGGCGTGC AAGATTCCGA ATACCGCAAG
5881 CGACAGGCGG ATCATCGTCG CGCTCCAGCG AAAGCGGTCC TCGCCGAAAA TGACCCAGAG
5941 CGCTGCCGGC ACCTGTCCTA CGAGTTGCAT GATAAAGAAG ACAGTCATAA GTGCGGCGAC
6001 GATAGTCATG CCGCGCGCCC ACCGGAAGGA GCTGACTGGG TTGAAGGCTC TCAAGGGCAT
6061 CCGTCCGATC ACGCTCTCCC TTATGCGACT CCTGCATTAG GAAGCAGCCC AGTAGTAGGT
6121 TGAGGCCGTT GAGCACCAGC GCCGCAAGGA ATGGTGCATG CAAGGAGATG GCGCCCAACA-

FIGURE 37C

98/240

6181 GTCCCCGGC CACGGGGCCT GCCACCATAC CCACGCCGAA ACAAGCGCTC ATGAGCCCGA
6241 AGTGGCGAGC CCGATCTTCC CCATCGGTGA TGTCGGCGAT ATAGGCGCCA GCAACCGCAC
6301 CTGTGGCGCC GGTGATGCCG GCCACGATGC GTCCGGCGTA GAGGATCGAG ATCT

FIGURE 37D

Figure 38A: pDEST18

FastBac Transfer Vector with p10
Baculovirus Promoter

1 gaagacctcg gccgtcgagg cgcttgccgg tgggtgetgac cccggatgaa gtggttcgca
cttctgggagc cggcagcgcc gcgaacggcc accacgactg gggcctactt caccaagcgt

61 tectcggttt tctggaagge gagcagctt tggtegecca ggactctagc tatagttcta
aggagccaaa agaccttccg ctgtagcaa acaagcgggt cctgagatcg atatcaagat

121 gtggttggtt acgtatcgag caagaaactg aaagccaaa tgcgttggag ptttctgac
caccaaccga tgcatagtct gttcttttat ttgctgggtt gcgaacctc agaacaagac

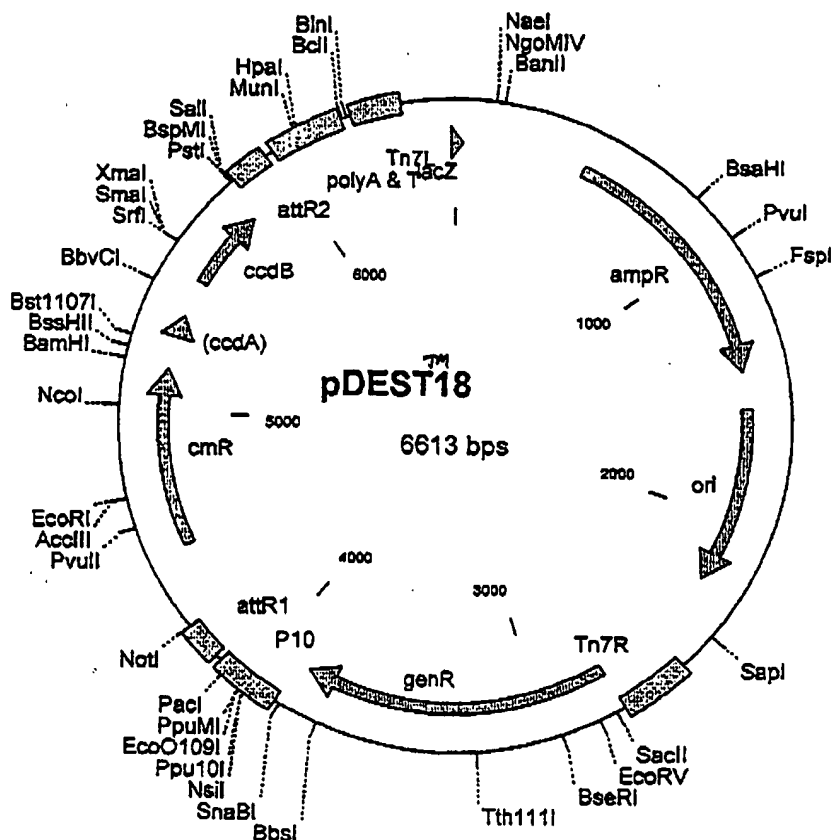
181 //tatcttaca agatctaga atctgcact actacacaa gggggacta caaatatag//
//ataaaactg ttctaaactt ttatgcctag tgaatctgt tccccctgat actttaatac//

241 //cattctgagg atgcccggac ctttatcca acccaacaca atatattata gtaaatatag// mRNA
//gaaaactcc tacggccctg gaaattaat tgggtctgt tatataaat caatttatc//

301 //atttatata caaatcattt gtatattaat caaatatata tactgtaaat tacattctat
//taataaata gtttagtata oataaatta attttatgat atgacattca atgtaaaata

361 ttacaatgag gatcatcaca agtttgtaca aaaaagctga acgagaaacg taaaatgata
aatgttactc ctagtagct tcaaacatgt tttctcgact tgctcttgc attttactat//

Int ↓ attRI



pDEST18 6613 bp

<u>Location (Base Nos.)</u>			<u>Gene Encoded</u>			
	474..1449		ampR			
	1590..2244		ori			
	2738..3850		genR			
	4251..4127		attR1			
	4501..5160		CmR			
	5280..5364		inactivated ccdA			
	5502..5807		ccdB			
	5848..5972		attR2			
	6595..25		lacZ			
1	GACGCGCCCT	GTAGCGGCGC	ATTAAGCGCG	GCGGGTGTGG	TGGTTACGCG	CAGCGTGACC
61	GCTACACTTG	CCAGCGCCCT	AGCGCCCGCT	CCTTTCGCTT	TCTTCCCTTC	CTTTCTCGCC
121	ACGTTTCGCCG	GCTTTCCTCCG	TCAAGCTCTA	AATCGGGGGC	TCCCTTTAGG	GTTCCGATTT
181	AGTGCTTTAC	GGCACCTCGA	CCCCAAAAAA	CTTGATTAGG	GTGATGGTTC	ACGTAGTGGG
241	CCATCGCCCT	GATAGACGGT	TTTTCGCCCT	TTGACGTTGG	AGTCCACGTT	CTTTAATAGT
301	GGACTCTTGT	TCCAAACTGG	AACAACACTC	AACCCTATCT	CGGTCTATTC	TTTTGATTTA
361	TAAGGGATTT	TGCCGATTTC	GGCCTATTGG	TTAAAAAATG	AGCTGATTTA	ACAAAAATTT
421	AACGCGAATT	TTAACAAAAT	ATTAACGTTT	ACAATTTTCAG	GTGGCACTTT	TCGGGGAAAT
481	GTGCGCGGAA	CCCCTATTTG	TTTATTTTTTC	TAAATACATT	CAAATATGTA	TCCGCTCATG
541	AGACAATAAC	CCTGATAAAT	GCTTCAATAA	TATTGAAAAA	GGAAGAGTAT	GAGTATTCAA
601	CATTTCCGTG	TCGCCCTTAT	TCCCTTTTTT	GCGGCATTTT	GCCTTCCTGT	TTTTGCTCAC
661	CCAGAAACGC	TGGTGAAAGT	AAAAGATGCT	GAAGATCAGT	TGGGTGCACG	AGTGGGTTAC
721	ATCGAACTGG	ATCTCAACAG	CGGTAAGATC	CTTGAGAGTT	TTGCCCCCGA	AGAACGTTTT
781	CCAATGATGA	GCACTTTTAA	AGTTCTGCTA	TGTGGCGCGG	TATTATCCCG	TATTGACGCC
841	GGGCAAGAGC	AACTCGGTCG	CCGCATACAC	TATTCTCAGA	ATGACTTGGT	TGAGTACTCA
901	CCAGTCACAG	AAAAGCATCT	TACGGATGGC	ATGACAGTAA	GAGAATTATG	CAGTGCTGCC
961	ATAACCATGA	GTGATAACAC	TGCGGCCAAC	TTACTTCTGA	CAACGATCGG	AGGACCGAAG
1021	GAGCTAACCG	CTTTTTTGCA	CAACATGGGG	GATCATGTAA	CTCGCCTTGA	TCGTTGGGAA
1081	CCGGAGCTGA	ATGAAGCCAT	ACCAAACGAC	GAGCGTGACA	CCACGATGCC	TGTAGCAATG
1141	GCAACAACGT	TGCGCAAACT	ATTAAGTGGC	GAAGTACTTA	CTCTAGCTTC	CCGGCAACAA
1201	TTAATAGACT	GGATGGAGGC	GGATAAAGTT	GCAGGACCAC	TTCTGCGCTC	GGCCCTTCCG
1261	GCTGGCTGGT	TTATTGCTGA	TAAATCTGGA	GCCGGTGAGC	GTGGGTCTCG	CGGTATCATT
1321	GCAGCACTGG	GGCCAGATGG	TAAGCCCTCC	CGTATCGTAG	TTATCTACAC	GACGGGGAGT
1381	CAGGCAACTA	TGGATGAACG	AAATAGACAG	ATCGCTGAGA	TAGGTGCCTC	ACTGATTAAG
1441	CATTGGTAAC	TGTCAGACCA	AGTTTACTCA	TATATACTTT	AGATTGATTT	AAAACCTCAT
1501	TTTTAATTTA	AAAGGATCTA	GGTGAAGATC	CTTTTTGATA	ATCTCATGAC	CAAAATCCCT
1561	TAACGTGAGT	TTTCGTTCCA	CTGAGCGTCA	GACCCCGTAG	AAAAGATCAA	AGGATCTTCT
1621	TGAGATCCCT	TTTTTCTGCG	CGTAATCTGC	TGCTTGCAAA	CAAAAAAACC	ACCGCTACCA
1681	GCGGTGGTTT	GTTTGCCGGA	TCAAGAGCTA	CCAAGTCTTT	TTCCGAAGGT	AACTGGCTTC
1741	AGCAGAGCGC	AGATACCAAA	TACTGTCCTT	CTAGTGTAGC	CGTAGTTAGG	CCACCACTTC
1801	AAGAAGTCTG	TAGCACCGCC	TACATACCTC	GCTCTGCTAA	TCCTGTTACC	AGTGGCTGCT
1861	GCCAGTGGCG	ATAAGTCGTG	TCTTACCGGG	TTGGAAGTCA	GACGATAGTT	ACCGGATAAG
1921	GCGCAGCGGT	CGGGCTGAAC	GGGGGGTTTC	TGCACACAGC	CCAGCTTGGG	GCGAACGACC
1981	TACACCGAAC	TGAGATACCT	ACAGCGTGAG	CATTGAGAAA	GCGCCACGCT	TCCCGAAGGG
2041	AGAAAGGCGG	ACAGGTATCC	GGTAAGCGGC	AGGGTTCGGAA	CAGGAGAGCG	CACGAGGGAG
2101	CTTCCAGGGG	GAAACGCCTG	GTATCTTTAT	AGTCCTGTCT	GGTTTCGCCA	CCTCTGACTT
2161	GAGCGTCGAT	TTTTGTGATG	CTCGTCAGGG	GGGCGGAGCC	TATGGAAAAA	CGCCAGCAAC
2221	GCGGCCTTTT	TACGGTTCCT	GGCCTTTTGC	TGGCCTTTTG	CTCACATGTT	CTTTCCTGCG
2281	TTATCCCTTG	ATTCTGTGGA	TAACCGTATT	ACCGCCTTTG	AGTGAGCTGA	TACCGCTCGC
2341	CGCAGCCGAA	CGACCGAGCG	CAGCGAGTCA	GTGAGCGTAG	AAGCGGAAGA	GCGCCTGATG
2401	CGGTATTTTC	TCCTTACGCA	TCTGTGCGGT	ATTTCACACC	GCAGACCAGC	CGCGTAACCT
2461	GGCAAAATCG	GTTACGGTTG	AGTAATAAAT	GGATGCCCTG	CGTAAGCGGG	TGTGGGCGGA-

FIGURE 3B

2521 CAATAAAGTC TTAAACTGAA CAAAATAGAT CTAAACTATG ACAATAAAGT CTTAAACTAG
2581 ACAGAATAGT TGTAACCTGA AATCAGTCCA GTTATGCTGT GAAAAAGCAT ACTGGAATTT
2641 TGTTATGGCT AAAGCAAACCT CTTCAATTTTC TGAAGTGCAA ATTGCCCGTC STATTAAAGA
2701 GGGGCGTGGC CAAGGGCATG GTAAAGACTA TATTCGCGGC GTTGTGACAA TTTACCGAAC
2761 AACTCCGCGG CCGGGAAGCC GATCTCGGCT TGAACGAATT GTTAGGTGGC SGTACTTGGG
2821 TCGATATCAA AGTGCATCAC TTCTTCCCGT ATGCCCAACT TTGTATAGAG AGCCACTGCG
2881 GGATCGTCAC CGTAATCTGC TTGCACGTAG ATCACATAAG CACCAAGCGC GTTGGCCTCA
2941 TGCTTGAGGA GATTGATGAG CGCGGTGGCA ATGCCCTGCC TCCGGTGCTC GCCGGAAGT
3001 GCGAGATCAT AGATATAGAT CTCACTACGC GGCTGCTCAA ACCTGGGCAG AACGTAAGCC
3061 GCGAGAGCGC CAACAACCGC TTCTTGCTCG AAGGCAGCAA GCGCGATGAA TGTCTTACTA
3121 CGGAGCAAGT TCCCGAGGTA ATCGGAGTCC GGCTGATGTT GGGAGTAGGT GGCTACGTCT
3181 CCGAACTCAC GACCGAAAAG ATCAAGAGCA GCCCGCATGG ATTGACTTG GTCAGG3CCG
3241 AGCCTACATG TGCGAATGAT GCCCATACTT GAGCCACCTA ACTTTGTTTT AGGGCGACTG
3301 CCTGCTGCG TAACATCGTT GCTGCTGCGT AACATCGTTG CTGCTCCATA ACATCAAACA
3361 TCGACCCACG GCGTAACGCG CTTGCTGCTT GGATGCCCGA GGCATAGACT GTACAAAAAA
3421 ACAGTGCATG CAAGCCATGA AAACCGCCAC TGCGCCGTTA CCACCGCTGC GTTCGGTCAA
3481 GGTCTCTGGAC CAGTTGCGTG AGCGCATACG CTACTTGCAT TACAGTTTAC GAACCGAACA
3541 GGCTTATGTC AACTGGGTTT GTGCCTTCAT CCGTTTCCAC GGTGTGCGTC ACCCGGCAAC
3601 CTTGGGCAGC AGCGAAGTCG AGGCATTTCT GTCCTGGCTG GCGAACGAGC GCAAGGTTTC
3661 GGTCCTCCAG CATCGTCAGG CATTGGCGGC CTTGCTGTTT TTCTACGGCA AGGTGCTGTG
3721 CACGGATCTG CCTGGGCTTC AGGAGATCGG AAGACCTCGG CCGTCGCGGC GCTTGGCGGT
3781 GGTGCTGACC CCGGATGAAG TGGTTCGCAT CCTCGGTTTT CTGGAAGGCG AGCATCSTTT
3841 GTTCGCCCCAG GACTCTAGCT ATAGTTCTAG TGGTTGGCTA CGTATCGAGC AAGAAAAATA
3901 AACGCCAAAC GCGTTGGAGT CTTGTGTGCT ATTTTACAA AGATTCAGAA ATACGCATCA
3961 CTTACAACAA GGGGGACTAT GAAATTATGC ATTTTGAGGA TGCCGGGACC TTTAATTCAA
4021 CCCAACACAA TATATTATAG TTAAATAAGA ATTATTTATC AAATCATTTG TATATTAAAT
4081 AAAATACTAT ACTGTAAATT ACATTTTATT TACAATGAGG ATCATCACAA GTTTGTACAA
4141 AAAAGCTGAA CGAGAAACGT AAAATGATAT AAATATCAAT ATATTAAAT AGATTTTGCA
4201 TAAAAAACAG ACTACATAAT ACTGTAAAC ACAACATATC CAGTCACTAT GCGCGCTGCT
4261 AAGTTGGCAG CATCACCCGA CGCACTTTGC GCCGAATAAA TACCTGTGAC GGAAGATCAC
4321 TTCGCAGAAT AAATAAATCC TGGTGTCCCT GTTGATACCG GGAAGCCCTG GGCCAACTTT
4381 TGGCGAAAAT GAGACGTTGA TCGGCACGTA AGAGGTTCCA ACTTTACCA TAATGAATA
4441 AGATCACTAC CGGGCGTATT TTTTGAGTTA TCGAGATTTT CAGGAGCTAA GGAAGCTAAA
4501 ATGGAGAAAA AAATCACTGG ATATACCACC GTTGATATAT CCCAATGGCA TCGTAAGAA
4561 CATTTTGAGG CATTTCAGTC AGTTGCTCAA TGTACCTATA ACCAGACCGT TCAGCTSGAT
4621 ATTACGGCCT TTTTAAAGAC CGTAAAGAAA AATAAGCACA AGTTTTATCC GGCCTTTATT
4681 CACATTTCTG CCCGCTGAT GAATGCTCAT CCGGAATTCC GTATGGCAAT GAAAGACGGT
4741 GAGCTGGTGA TATGGGATAG TGTTACCCCT TGTACACCG TTTTCCATGA GCAAACGAA
4801 ACGTTTTTCAT CGCTCTGGAG TGAATACCAC GACGATTTCC GGCAGTTTCT ACACATATAT
4861 TCGCAAGATG TGGCGTGTTA CCGTGAAAAC CTGGCCTATT TCCCTAAAGG GTTTATTGAG
4921 AATATGTTTT TCGTCTCAGC CAATCCCTGG GTGAGTTTCA CCAGTTTGA TTTAAACGTG
4981 GCCAATATGG ACAACTTCTT CGCCCCCGTT TTCACCATGG GCAAATATTA TACGCAAGGC
5041 GACAAGGTGC TGATGCCGCT GCGGATTGAG GTTCATCATG CCGTCTGTGA TGGCTTCCAT
5101 GTCGGCAGAA TGCTTAATGA ATTACAACAG TACTGCGATG AGTGGCAGG CGGGGCTAA
5161 ACGCGTGGAT CCGGCTTACT AAAAGCCAGA TAACAGTATG CGTATTTGCG CGCTGATTTT
5221 TGCGGTATAA GAATATATAC TGATATGTAT ACCCGAAGTA TGTCAAAAAG AGGTGTGCTA
5281 TGAAGCAGCG TATTACAGTG ACAGTTGACA GCGACAGCTA TCAGTTGCTC AAGGCATATA
5341 TGATGTCAAT ATCTCCGGTC TGGTAAGCAC AACCATGCAG AATGAAGCCC GTCGCTGCG
5401 TGCCGAACGC TGGAAAGCGG AAAATCAGGA AGGGATGGCT GAGGTGCGCC GGTTTATTGA
5461 AATGAACGCG TCTTTTGTG ACGAGAACAG GGAAGGTGTA AATGCAGTTT AAGGTTTACA
5521 CCTATAAAAG AGAGAGCCGT TATCGTCTGT TTGTGGATGT ACAGAGTGAT ATTATTGACA
5581 CGCCCGGGCG ACGGATGGTG ATCCCGCTGT CCAAGTGACG TCTGCTGTCA GATAAAGTCT
5641 CCGGTGAAC TTAACCGGTG GTGCATATCG GGGATGAAAG CTGGCGCATG ATGACCACCG
5701 ATATGGCCAG TGTGCCGGTC TCCGTTATCG GGGGAAGAAGT GGCTGATCTC AGCCACCGCG
5761 AAAATGACAT CAAAAACGCC ATTAACCTGA TGTCTGGGG AATATAAATG TCAGGCTCCC
5821 TTATACACAG CCAGTCTGCA GGTCGACCAT AGTGACTGGA TATGTTGTGT TTTACAGTAT
5881 TATGTAGTCT GTTTTTTATG CAAAATCTAA TTTAATATAT TGATATTTAT ATCATTTTAC
5941 GTTCTCGTT CAGCTTTCTT GTACAAAGTG GTGATAGCTT GTCGAGAAGT ACTAGAGGAT

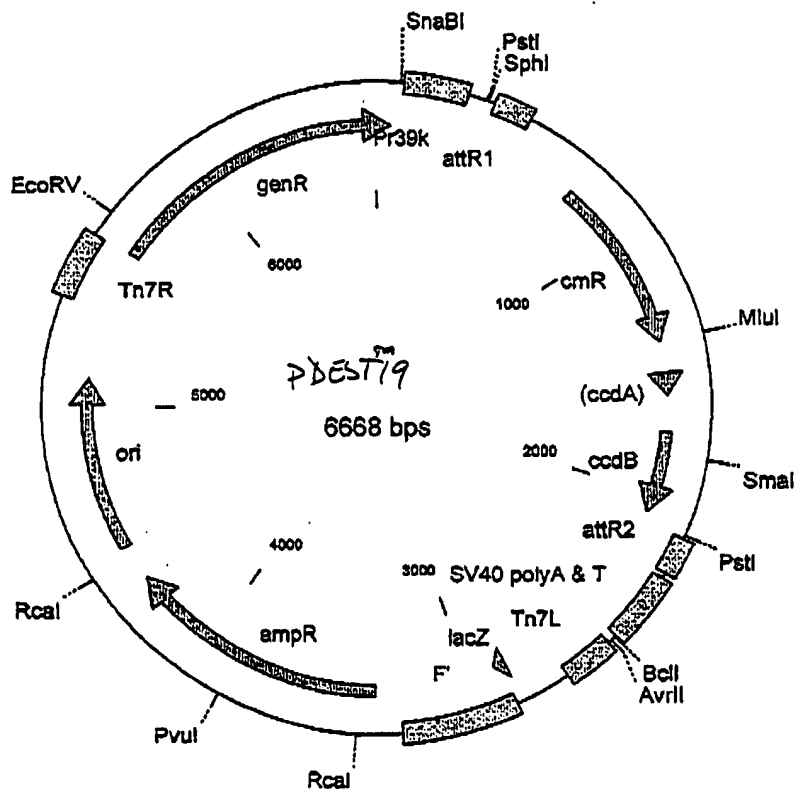
102/240

6001 CATAATCAGC CATACCACAT TTGTAGAGGT TTTACTTGCT TTAAAAAACC TCCCACACCT
6061 CCCCCTGAAC CTGAAACATA AAATGAATGC AATTGTTGTT GTTAACTTGT TTATTGCAGC
6121 TTATAATGGT TACAAATAAA GCAATAGCAT CACAAATTTT ACAAATAAAG CATTTTTTTTC
6181 ACTGCATTCT AGTTGTGGTT TGTCCAAACT CATCAATGTA TCTTATCATG TCTGGATCTG
6241 ATCACTGCTT GAGCCTAGGA GATCCGAACC AGATAAGTGA AATCTAGTTC CAAACTATTT
6301 TGTCATTTTT AATTTTCGTA TTAGCTTACG ACGCTACACC CAGTTCCCAT CTATTTTGTC
6361 ACTCTTCCCT AAATAATCCT TAAAAACTCC ATTTCCACCC CTCCAGTTC CCAACTATTT
6421 TGTCCGCCCA CAGCGGGGCA TTTTCTTCC TGTTATGTTT TTAATCAAAC ATCCTGCCAA
6481 CTCCATGTGA CAAACCGTCA TCTTCGGCTA CTTTTTCTCT GTCACAGAAT GAAAATTTTT
6541 CTGTCATCTC TTCGTTATTA ATGTTTGTA TTAGCTGAAT ATCAACGCTT ATTTGCAGCC
6601 TGAATGGCGA ATG

FIGURE 38D

103/240

1 ggtgacgccc tcatttttcc attgtaacgt aaatggcaac ttgtagatga acgcgctgtc
 ccactgccgc agtagaaaagg taacattgca ttaccggtg aacatctact tgcgcgacag
 61 aaaaaaccgg ccagttttctt ccacaaactc gcgcacggct gtctcgtaaa cttttgcgtc
 ttttttggcc ggtcaaagaa ggtgtttgag cgcgtgccga cagagcattt gaaaaacgag
 121 // gcaacaatcg cgatgacctc gtggtatgga aattttttct aaaaaagtgt cgttcattgtc //
 // cgttggttagc gctactggag caccatacct ttaaaaaaga ttttttcaca gcaagtacag //
 181 // ggccggggcg ttcgcgctcc ggtacgcgcg acgggcacac agcaggacag ctttgctcgg //
 // ccgcgcgcgc aagcgcgagg ccatgcgcgc tgcgcgtgtg tcgtccctgc ggaacaggcc //
 241 etcgattatc ataaacaatc ctgcaggcat gcaagctgga tcatacaag ttgtacaaa
 gagctaatag tatttgtag gacgtccgta cgttcgacct agtagtcttc aaacatgctt
 Int V



104/240

pDEST19 6668 bp (rotated to position 1000)

<u>Location (Base Nos.)</u>		<u>Gene Encoded</u>
515..391		attR1
765..1424		CmR
1544..1628		inactivated ccdA
1766..2071		ccdB
2112..2236		attR2
2852..2895		lacZ
3344..4319		ampR
4460..5114		ori
5608..52		genR
1	AGTGGTTCGC ATCCTCGGTT TTCTGGAAGG CGAGCATCGT TTGTTTCGCC AGGACTCTAG	
61	CTATAGTTCT AGTGGTTGGC TACGTATATC AAATACTTGT AGGTGACGCC GTCATCTTTC	
121	CATTGTAACG TAAATGGCAA CTTGTAGATG AACGCGCTGT CAAAAAACCG GCCAGTTTCT	
181	TCCACAAACT CGCGCACGGC TGTCTCGTAA ACTTTTGCCT CGCAACAATC GCGATGACCT	
241	CGTGGTATGG AAATTTTTTC TAAAAAAGTG TCGTTCATGT CGGCGGCGGG CGCGTTCGCG	
301	CTCCGGTACG CGCGACGGGC ACACAGCAGG ACAGCCTTGT CCGGCTCGAT TATCATAAAC	
361	AATCCTGCAG GCATGCAAGC TCGGATCATC ACAAGTTTGT ACAAAAAAGC TGAACGAGAA	
421	ACGTAAAATG ATATAAATAT CAATATATTA AATTAGATTT TGCATAAAAA ACAGACTACA	
481	TAATACTGTA AAACACAACA TATCCAGTCA CTATGGCGGC CGCTAAGTTG GCAGCATCAC	
541	CCGACGCACT TTGCGCCGAA TAAATACCTG TGACGGAAGA TCACTTCGCA GAATAAATAA	
601	ATCCTGGTGT CCCTGTTGAT ACCGGGAAGC CCTGGGCCAA CTTTGGCGA AAATGAGACG	
661	TTGATCGGCA CGTAAGAGGT TCCAACCTTC ACCATAATGA AATAAGATCA CTACCGGGCG	
721	TATTTTTTGA GTTATCGAGA TTTTCAGGAG CTAAGGAAGC TAAAATGGAG AAAAAATCA	
781	CTGGATATAC CACCGTTGAT ATATCCCAAT GGCAATCGTAA AGAACATTTT GAGGCATTTC	
841	AGTCAGTTGC TCAATGTACC TATAACCAGA CCGTTCAGCT GGATATTACG GCCTTTTAA	
901	AGACCGTAAA GAAAAATAAG CACAAGTTTT ATCCGGCCTT TATTCACATT CTGCCCCGCC	
961	TGATGAATGC TCATCCGGAA TTCCGTATGG CAATGAAAGA CGGTGAGCTG GTGATATGGG	
1021	ATAGTGTTCA CCCTTGTTAC ACCGTTTTCC ATGAGCAAAC TGAAACGTTT TCATCGCTCT	
1081	GGAGTGAATA CCACGACGAT TTCCGGCAGT TTCTACACAT ATATTCGCAA GATGTGGCGT	
1141	GTTACGGTGA AAACCTGGCC TATTTCCCTA AAGGGTTTAT TGAGAATATG TTTTTCGTCT	
1201	CAGCCAATCC CTGGGTGAGT TTCACCACTT TTGATTTAAA CGTGGCCAAT ATGGACAAC	
1261	TCTTCGCCCC CGTTTTACAC ATGGGCAAAT ATTATACGCA AGGCGACAAG GTGCTGATGC	
1321	CGCTGGCGAT TCAGGTTTCAT CATGCCGCTT GTGATGGCTT CCATGTCGGC AGAATGCTTA	
1381	ATGAATTACA ACAGTACTGC GATGAGTGCC AGGGCGGGGC GTAAACGCGT GGATCCGGCT	
1441	TACTAAAAAGC CAGATAACAG TATGCGTATT TGCGCGCTGA TTTTTCGGT ATAAGAATAT	
1501	ATACTGATAT GTATACCCGA AGTATGTCAA AAAGAGGTGT GCTATGAAGC AGCGTATTAC	
1561	AGTGACAGTT GACAGCGACA GCTATCAGTT GCTCAAGGCA TATATGATGT CAATATCTCC	
1621	GGTCTGGTAA GCACAACCAT GCAGAATGAA GCGCGTCGTC TCGTGCCGA ACGCTGGA	
1681	GCGGAAAATC AGGAAGGGAT GGCTGAGGTC GCGCGGTTTA TTGAAATGAA CGGCTCTTTT	
1741	GCTGACGAGA ACAGGGACTG GTGAAATGCA GTTTAAGGTT TACACCTATA AAAGAGAGAG	
1801	CCGTTATCGT CTGTTGTGG ATGTACAGAG TGATATTATT GACACGCCCC GGCGACGGAT	
1861	GGTGATCCCC CTGGCCAGTG CACGTCTGCT GTCAGATAAA GTCTCCCGTG AACTTTACCC	
1921	GGTGGTGCAT ATCGGGGATG AAAGCTGGCG CATGATGACC ACCGATATGG CCAAGTGGCC	
1981	GGTCTCCGTT ATCGGGGAAG AAGTGGCTGA TCTCAGCCAC CGCGAAAATG ACATCAAAAA	
2041	CGCCATTAACT CTGATGTTCT GGGGAATATA AATGTCAGGC TCCCTTATAC ACAGCCAGTC	
2101	TGCAGTTCGA CCATAGTGAC TGGATATGTT GTGTTTTACA GTATTATGTA GTCTGTTTTT	
2161	TATGCAAAAT CTAATTTAAT ATATTGATAT TTATATCATT TTACGTTTCT CGTTCAGCTT	
2221	TCTTGTAACA AGTGGTGATC GAGAAGTACT AGAGGATCAT AATCAGCCAT ACCACATTTG	
2281	TAGAGGTTTT ACTTGCTTTA AAAAACTCC CACACCTCCC CCTGAACCTG AAACATAAAA	
2341	TGAATGCAAT TGTGTGTTGTT AACTTGTTTA TTGCAGCTTA TAATGGTTAC AAATAAGCA	
2401	ATAGCATCAC AAATTTTACA AATAAAGCAT TTTTTTCACT GCATTCTAGT TGTGGTTTGT	
2461	CCAACTCAT CAATGTATCT TATCATGTCT GGATCTGATC ACTGCTTGAG CCTAGGAGAT	
2521	CCGAACCAGA TAAGTGAAAT CTAGTTCCAA ACTATTTTGT CATTTTTAAT TTTTCGTATTA	
2581	GCTTACGACG CTACACCCAG TTCCCATCTA TTTTGTCACT CTTCCCTAAA TAATCCTTAA-	

FIGURE 39B

105/240

2641 AAACCTCCATT TCCACCCCTC CCAGTTCCCA ACTATTTTGT CCGCCACAG CGGGGCATTT
2701 TTCTTCCTGT TATGTTTTTA ATCAAACATC CTGCCAACTC CATGTGACAA ACCGTCATCT
2761 TCGGCTACTT TTTCTCTGTC ACAGAATGAA AATTTTTCTG TCATCTCTTC GTTATTAATG
2821 TTTGTAATTG ACTGAATATC AACGCTTATT TGCAGCCTGA ATGGCGAATG GACGCGCCCT
2881 GTAGCGGCGC ATTAAGCGCG GCGGGTGTGG TGGTTACGCG CAGCGTGACC GCTACACTTG
2941 CCAGCGCCCT AGCGCCCGCT CCTTTCGCTT TCTTCCCTTC CTTTCTCGCC ACGTTCGCGG
3001 GCTTTCCTCCG TCAAGCTCTA AATCGGGGGC TCCCTTTAGG GTTCCGATTT AGTGCTTTAC
3061 GGCACCTCGA CCCCCAAAAA CTTGATTAGG GTGATGGTTC ACGTAGTGGG CCATCGCCCT
3121 GATAGACGGT TTTTCGCCCT TTGACGTTGG AGTCCACGTT CTTTAATAGT GGACTCTTGT
3181 TCCAAACTGG AACAACTC AACCCCTATCT CGGTCTATTC TTTTGATTTA TAAGGGATTT
3241 TGCCGATTTC GGCCTATTGG TTAATAAATG AGCTGATTTA AAAAAATTT AACCGAATT
3301 TTAACAAAAT ATTAACGTTT ACAATTTTCA GTGGCACTTT TCGGGGAAAT GTGCGCGGAA
3361 CCCCTATTTG TTTATTTTTT TAAATACATT CAAATATGTA TCCGCTCATG AGACAATAAC
3421 CCTGATAAAT GCTTCAATAA TATTGAAAAA GGAAGAGTAT GAGTATTCAT CATTTCCGTG
3481 TCGCCCTTAT TCCCTTTTTT GCGGCATTTT GCCTTCCTGT TTTTGCTCAC CCAGAAACGC
3541 TGGTGAAAGT AAAAGATGCT GAAGATCAGT TGGGTGCACG AGTGGGTAC ATCGAACTGG
3601 ATCTCAACAG CGGTAAGATC CTTGAGAGTT TTCGCCCCGA AGAACGTTTT CCAATGATGA
3661 GCACTTTTAA AGTTCCTGCTA TGTGGCGCGG TATTATCCCG TATTGACGCC GGGCAAGAGC
3721 AACTCGGTCG CCGCATACAC TATTCTCAGA ATGACTTGGT TGAGTACTCA CCAGTCACAG
3781 AAAAGCATCT TACGGATGGC ATGACAGTAA GAGAATTATG CAGTGCTGCC ATAACCATGA
3841 GTGATAACAC TGCGGCCAAC TTACTTCTGA CAACGATCGG AGGACCGAAG GAGCTAACCG
3901 CTTTTTTGCA CAACATGGGG GATCATGTAA CTCGCCTTGA TCGTTGGGAA CCGGAGCTGA
3961 ATGAAGCCAT ACCAAACGAC GAGCGTGACA CCACGATGCC TGTAAGCAATG GCAACAACGT
4021 TGCGCAAACT ATTAACGTCG GAACTACTTA CTCTAGCTTC CCGGCAACAA TTAATAGACT
4081 GGATGGAGGC GGATAAAGTT GCAGGACCAC TTCTGCGCTC GGCCCTTCCG GCTGGCTGGT
4141 TTATTGCTGA TAAATCTGGA GCCGGTGAGC GTGGGTCTCG CGGTATCATT GCAGCACTGG
4201 GGCCAGATGG TAAGCCCTCC CGTATCGTAG TTATCTACAC GACGGGGAGT CAGGCAACTA
4261 TGGATGAACG AAATAGACAG ATCGCTGAGA TAGGTGCCTC ACTGATTAAG CATTGGTAAAC
4321 TGTCAGACCA AGTTTACTCA TATATACTTT AGATTGATTT AAAAATTTCAT TTTTAATTTA
4381 AAAGGATCTA GGTGAAGATC CTTTTTGATA ATCTCATGAC CAAAATCCCT TAACGTGAGT
4441 TTTCTGTTCCA CTGAGCGTCA GACCCCGTAG AAAAGATCAA AGGATCTTCT TGAGATCCTT
4501 TTTTCTGCG CGTAATCTGC TGCTTGCAAA CAAAAAACC ACCGCTACCA GCGGTGGTTT
4561 GTTTGCCGGA TCAAGAGCTA CCAACTCTTT TTCCGAAGGT AACTGGCTTC AGCAGAGCGC
4621 AGATACCAAA TACTGTCTT CTAGTGATAG CGTAGTTAGG CCACCACTTC AAGAACTCTG
4681 TAGCACCGCC TACATACCTC GCTCTGCTAA TCCTGTTACC AGTGGCTGCT GCCAGTGGCG
4741 ATAAGTCGTG TCTTACCGGG TTGACTCAA GACGATAGTT ACCGGATAAG GCGCAGCGGT
4801 CGGGCTGAAC GGGGGGTTTC TGCACACAGC CCAGCTTGA GCGAACGACC TACACCGAAC
4861 TGAGATACCT ACAGCGTGAG CATTGAGAAA GCGCCACGCT TCCGAAGGG AGAAAGGCGG
4921 ACAGGTATCC GGTAAGCGGC AGGGTCGGAA CAGGAGAGCG CACGAGGGAG CTTCCAGGGG
4981 GAAACGCGCTG GTATCTTTAT AGTCCTGTG GGTTCGCCA CCTCTGACTT GAGCGTCGAT
5041 TTTTGTGATG CTCGTACAGG GGGCGGAGCC TATGAAAAA CGCCAGCAAC GCGGCCTTTT
5101 TACGGTTCCT GGCCTTTTGC TGGCCTTTTG CTCACATGTT CTTTCTGCG TTATCCCTTG
5161 ATTCTGTGGA TAACCGTATT ACCGCCTTTG AGTGAGCTGA TACCGCTCGC CGCAGCCGAA
5221 CGACCGAGCG CAGCGAGTCA GTGAGCGAGG AAGCGGAAGA GCGCCTGATG CGGTATTTTC
5281 TCCTTACGCA TCTGTGCGGT ATTTACACAC GCAGACCAGC CGGTAAACCT GGCAAAATCG
5341 GTTACGGTTG AGTAATAAAT GGATGCCCTG CGTAAGCGGG TGTGGGCGGA CAATAAAGTC
5401 TTAAGCTGAA CAAAATAGAT CTAAACTATG ACAATAAAGT CTTAAACTAG ACAGAATAGT
5461 TGTAAGCTGA AATCAGTCCA GTTATGCTGT GAAAAAGCAT ACTGGACTTT TGTTATGGCT
5521 AAAGCAAACT CTTCAATTTT TGAAGTGCAA ATTGCCCGTC GTATTAAAGA GGGCGGTGGC
5581 CAAGGCGATG GTAAAGACTA TATTGCGCGC GTTGTGACAA TTTACCGAAC AACTCCGCGG
5641 CCGGGAAGCC GATCTCGGCT TGAACGAATT GTTAGGTGGC GGTACTTGGG TCGATATCAA
5701 AGTGATCAC TTCTTCCCGT ATGCCCAACT TTGTATAGAG AGCCACTGCG GGATCGTCAC
5761 CGTAATCTGC TTGCACGTAG ATCACATAAG CACCAAGCGC GTTGGCCTCA TGCTTGAGGA
5821 GATTGATGAG CGCGGTGGCA ATGCCCTGCC TCCGGTGCTC GCCGGAGACT GCGAGATCAT
5881 AGATATAGAT CTCACTACGC GGCTGCTCAA ACCTGGGCAG AACGTAAGCC GCGAGAGCGC
5941 CAACAACCGC TTCTTGGTGC AAGGCAGCAA GCGCGATGAA TGTCTTACTA CGGAGCAAGT
6001 TCCCAGGTA ATCGAGTCC GGCTGATGTT GGGAGTAGGT GGCTACGTCT CCGAACTCAC
6061 GACCGAAAAG ATCAAGAGCA GCCCGCATGG ATTTGACTTG GTCAGGGCCG AGCCTACATG-

FIGURE 39C

106/240

6121 TGCGAATGAT GCCCATACTT GAGCCACCTA ACTTTGTTTT AGGGCGACTG CCCTGCTGCG
6181 TAACATCGTT GCTGCTGCGT AACATCGTTG CTGCTCCATA ACATCAAACA TCGACCCACG
6241 GCGTAACGCG CTTGCTGCTT GGATGCCCGA GGCATAGACT GTACAAAAAA ACAGTCATAA
6301 CAAGCCATGA AAACCGCCAC TCGCGCGTTA CCACCGCTGC GTTCGGTCAA GGTTCTGGAC
6361 CAGTTGCGTG AGCGCATACG CTAATTGCAT TACAGTTTAC GAACCGAACA GGCTTATGTC
6421 AACTGGGTTT GTGCCTTCAT CCGTTTCCAC GGTGTGCGTC ACCCGGCAAC CTTGGGCAGC
6481 AGCGAAGTCG AGGCATTTCT GTCCTGGCTG GCGAACGAGC GCAAGGTTTC GGTCTCCACG
6541 CATCGTCAGG CATTGGCGGC CTTGCTGTTC TTCTACGGCA AGGTGCTGTG CACGGATCTG
6601 CCCTGGCTTC AGGAGATCGG AAGACCTCGG CCGTCGCGGC GCTTGCCGGT GGTGCTGACC
6661 CCGGATGA

FIGURE 39A

107/240

Figure 40A: pDEST20 Glutathione-S-transferase Fusion with Polyhedron Promoter for Baculovirus Expression

430 ggc tac gta tac tcc gga ata tta ata gat cat gga gat aat taa aat gat
ccg atg cat atg agg cct tat aat tat cta gta cct cta tta att tta cta

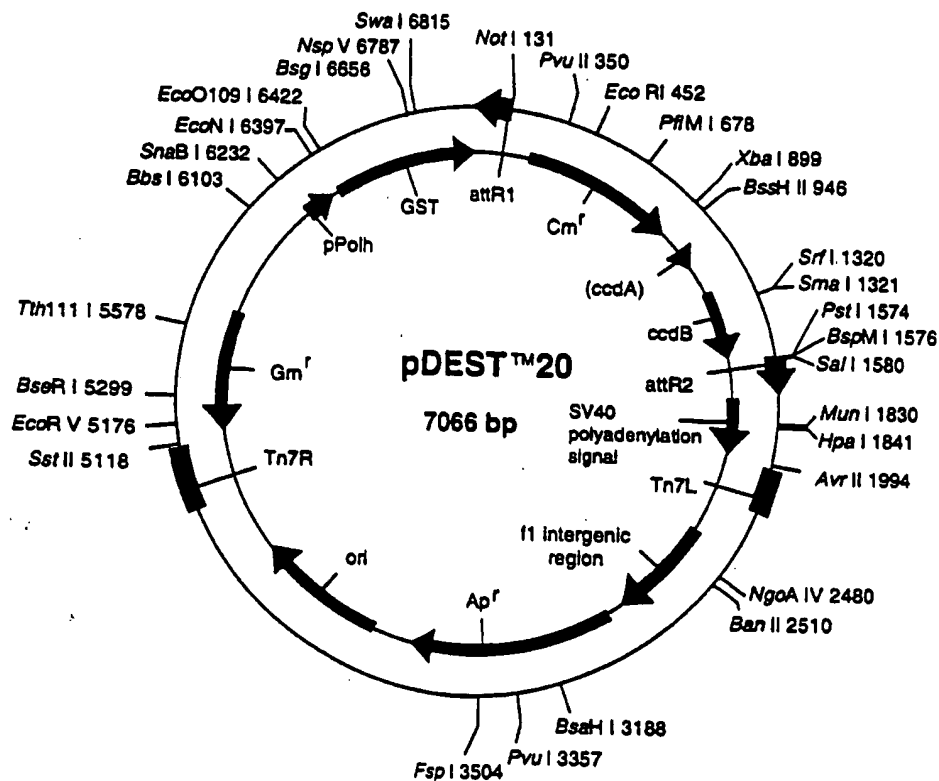
481 aac cat ctc gca aat aaa taa gta ttt tac tgt ttt cgt aac agt ttt gta
ttg gta gag cgt tta ttt att cat aaa atg aca aaa gca ttg tca aaa cat

532 ata aaa aaa cct ata aat att ccg gat tat tca tac cgt ccc acc atc ggg
tat ttt ttt gga tat tta taa ggc cta ata agt atg gca ggg tgg tag ccc

Start Transl. → A P I - - - GST - -
583 cgc gga tcc atg gcc cct ata cta ggt tat tgg aaa att aag ggc ctt gtg
gcg cct agg tac cgg gga tat gat cca ata acc ttt taa ttc ccg gaa cac //

1246 S D L V P R H N Q T S L Y K R A
tcg gat ctg gtt ccg cgt cat aat caa aca agt tgg tac aaa aaa gct gaa
agc cta gac caa ggc gca gta tta gtt tgc tca aac atg ttt ttt cga ctt

1297 cga gaa acg taa aat gat ata aat atc aat ata tta aat tag at
gct ctt tgc att tta cta tat tta tag tta tat aat tta atc ta



108/240

pDEST20 7066 bp (rotated to position 5800)

<u>Location (Base Nos.)</u>	<u>Gene Encoded</u>
592..1263	GST
1397..1273	attR1
1506..2165	CmR
2285..2369	inactivated ccdA
2507..2812	ccdB
2853..2977	attR2
4214..5064	ampR
5263..5843	ori

```

1  CCACTGCGCC GTTACCACCG CTGCGTTCCG TCAAGGTTCT GGACCAAGTTG CGTGAGCGCA
61 TACGCTACTT GCATTACAGT TTACGAACCG AACAGGCTTA TGTCAACTGG GTTCGTGCCT
121 TCATCCGTTT CCACGGTGTG CGTCACCCGG CAACCTTGGG CAGCAGCGAA GTCGAGGCAT
181 TTCTGTCTCG GCTGGCGAAC GAGCGCAAGG TTTCGGTCTC CACGCATCGT CAGGCATTGG
241 CGGCCTTGCT GTTCTTCTAC GGCAAGGTGC TGTGCACGGA TCTGCCCTGG CTTCAGGAGA
301 TCGGAAGACC TCGGCCGTGC CGGCGCTTGC CGGTGGTGCT GACCCCGGAT GAAAGTGGTTC
361 GCATCCTCGG TTTTCTGGAA GGCGAGCATC GTTTGTTTCG CCAGGACTCT AGCTATAGTT
421 CTAGTGGTTG GCTACGTATA CTCCGGAATA TTAATAGATC ATGGAGATAA TAAAATGAT
481 AACCATCTCG CAAATAAATA AGTATTTTAC TGTTTTCGTA ACAGTTTTGT AATAAAAAAA
541 CCTATAAATA TTCCGGATTA TTCATACCGT CCCACCATCG GGCGCGGATC CATGGCCCCCT
601 ATACTAGGTT ATTGGAAAAT TAAGGGCCCT GTGCAACCCA CTCGACTTCT TTTGGAATAT
661 CTTGAAGAAA AATATGAAGA GCATTTGTAT GAGCGCGATG AAGGTGATAA ATGGCGAAAC
721 AAAAAAGTTT AATTGGGTTT GGAGTTTCCC AATCTTCCTT ATTATATTGA TGGTGATGTT
781 AAATTAACAC AGTCTATGGC CATCATACGT TATATAGCTG ACAAGCACAA CATGTTGGGT
841 GGTTGTCCAA AAGAGCGTGC AGAGATTTCA ATGCTTGAAG GAGCGGTTTT GGATATTAGA
901 TACGGTGTTC CGAGAATTGC ATATAGTAAA GACTTTGAAA CTCTCAAAGT TGATTTTCTT
961 AGCAAGCTAC CTGAAATGCT GAAAATGTTT GAAGATCGTT TATGTCATAA AACATATTTA
1021 AATGGTGATC ATGTAACCCA TCCTGACTTC ATGTTGTATG ACGCTCTTGA TGTGTTTTTA
1081 TACATGGACC CAATGTGCCT GGATGCGTTC CCAAAATTAG TTTGTTTTAA AAAACGTATT
1141 GAAGACATCC CACAAATTGA TAAGTACTTG AAATCCAGCA AGTATATAGC ATGGCCCTTG
1201 CAGGGCTGGC AAGCCACGTT TGGTGGTGGC GACCATCCTC CAAAATCGGA TCTGGTTCGG
1261 CGTCATAATC AAACAAGTTT GTACAAAAAA GCTGAACGAG AAACGTAAAA TGATATAAAT
1321 ATCAATATAT TAAATTAGAT TTTGCATAAA AAACAGACTA CATAATACTG TAAACACAA
1381 CATATCCAGT CACTATGGCG GCCGCATTAG GCACCCAGG CTTTACACTT TATGCTTCCG
1441 GCTCGTATGT TGTGTGGATT TTGAGTTAGG ATCCGGCGAG ATTTTCAGGA GCTAAGGAAG
1501 CTAATAATGA GAAAAAATC ACTGGATATA CCACCGTTGA TATATCCCAA TGGCATCGTA
1561 AAGAACATTT TGAGGCATTT CAGTCAGTTG CTCAATGTAC CTATAACCAG ACCGTTACAG
1621 TGGATATTAC GGCCTTTTTA AAGACCGTAA AGAAAAATAA GCACAAGTTT TATCCGGCCT
1681 TTATTCACAT TCTTGCCCGC CTGATGAATG CTCATCCGGA ATTCCGTATG GCAATGAAAG
1741 ACGGTGAGCT GGTGATATGG GATAGTGTTC ACCCTTGTTA CACCGTTTTT CATGAGCAAA
1801 CTGAAACGTT TTCATCGCTC TGGAGTGAAT ACCACGACGA TTTCCGGCAG TTTCTACACA
1861 TATATTCGCA AGATGTGGCG GTTACGGTG AAAACCTGGC CTATTTCCCT AAAGGGTTTA
1921 TTGAGAATAT GTTTTTCGTC TCAGCCAATC CCTGGGTGAG TTTACCAGT TTTGATTTAA
1981 ACGTGGCCAA TATGGACAAC TTCTTCGCCC CCGTTTTTAC CATGGGCAAA TATTATACGC
2041 AAGGCGACAA GGTGCTGATG CCGCTGGCGA TTCAGGTTCA TCATGCCGTC TGTGATGGCT
2101 TCCATGTCGG CAGAATGCTT AATGAATTAC AACAGTACTG CGATGAGTGG CAGGGCGGGG
2161 CGTAATCTAG AGGATCCGGC TTAATAAAG CCAGATAACA GTATGCGTAT TTGCGCGCTG
2221 ATTTTTGCGG TATAAGAATA TATACTGATA TGTATACCCG AAGTATGTCA AAAAGAGGTG
2281 TGCTATGAAG CAGCGTATTA CAGTGACAGT TGACAGCGAC AGCTATCAGT TGCTCAAGGC
2341 ATATATGATG TCAATATCTC CGGTCTGGTA AGCACAACCA TGCAGAATGA AGCCCCGCTG
2401 CTGCGTGCCG AACGCTGGAA AGCGGAAATC CAGGAAGGGA TGGCTGAGGT CGCCCCGTTT
2461 ATTGAAATGA ACGGCTCTTT TGCTGACGAG AACAGGGACT GGTGAAATGC AGTTTAAGGT
2521 TTACACCTAT AAAAGAGAGA GCCGTTATCG TCTGTTTGTG GATGTACAGA GTGATATTAT
2581 TGACACGCCC GGGCGACGGA TGGTGATCCC CCTGGCCAGT GCACGTCTGC TGTGAGATAA
2641 AGTCTCCCGT GAACCTTACC CGGTGGTGCA TATCGGGGAT GAAAGCTGGC GCATGATGAC

```

Figure 40B

2701 CACCGATATG GCCAGTGTGC CGGTCTCCGT TATCGGGGAA GAAGTGGCTG ATCTCAGCCA
2761 CCGCGAAAAT GACATCAAAA ACGCCATTAA CCTGATGTC TGGGGAATAT AAATGTCAGG
2821 CTCCCTTATA CACAGCCAGT CTGCAGGTCG ACCATAGTGA CTGGATATGT TGTGTTTTAC
2881 AGTATTATGT AGTCTGTTTT TTATGCAAAA TCTAATTTAA TATATTGATA TTTATATCAT
2941 TTTACGTTTC TCGTTCAGCT TTCTTGTAACA AAGTGGTTTG ATAGCTTGTC GAGAAGTACT
3001 AGAGGATCAT AATCAGCCAT ACCACATTTG TAGAGGTTTT ACTTGCTTTA AAAAACCTCC
3061 CACACCTCCC CCTGAACCTG AAACATAAAA TGAATGCAAT TGTGTTGTT AACTTGTTTA
3121 TTGCAGCTTA TAATGGTTAC AAATAAAGCA ATAGCATCAC AAATTCACA AATAAAGCAT
3181 TTTTTTCACT GCATTCTAGT TGTGGTTTGT CCAAACCTCAT CAATGTATCT TATCATGTCT
3241 GGATCTGATC ACTGCTTGAG CCTAGGAGAT CCGAACCAGA TAAGTGAAAT CTAGTTCCAA
3301 ACTATTTTGT CATTTTAAAT TTTCGTATTA GCTTACGACG CTACACCCAG TTCCCATCTA
3361 TTTTGTCACT CTTCCCTAAA TAATCCTTAA AAACCTCCATT TCCACCCCTC CCAGTTCCCA
3421 ACTATTTTGT CCGCCACAG CGGGGCATTT TTCTTCCTGT TATGTTTTTA ATCAAACATC
3481 CTGCCAACTC CATGTGACAA ACCGTCATCT TCGGCTACTT TTTCTCTGTC ACAGAATGAA
3541 AATTTTTCTG TCATCTCTTC GTTATTAATG TTTGTAATTG ACTGAATATC AACGCTTATT
3601 TGCAGCCTGA ATGGCGAATG GACGCGCCCT GTAGCGGCGC ATTAAGCGCG GCGGGTGTGG
3661 TGGTTACGCG CAGCGTGACC GCTACACTTG CCAGCGCCCT AGCGCCCGCT CCTTTCGCTT
3721 TCTTCCCTTC CTTTCTCGCC ACGTTCGCGG GCTTTCCTCG TCAAGCTCTA AATCGGGGGC
3781 TCCCTTTAGG GTTCCGATTT AGTGCTTTAC GGCACCTCGA CCCCCAAAAA CTTGATTAGG
3841 GTGATGGTTC ACGTAGTGGG CCATCGCCCT GATAGACGGT TTTTCGCCCT TTGACGTGG
3901 AGTCCACGTT CTTTAATAGT GGACTCTTGT TCCAAACTGG AACAACACTC AACCCTATCT
3961 CGGTCTATTTC TTTTGATTTA TAAGGGATTT TGCCGATTTT GGCCTATTGG TTAATAAATG
4021 AGCTGATTTA ACAAAAATTT AACCGGAATT TTAACAAAAT ATTAACGTTT ACAATTTAG
4081 GTGGCACTTT TCGGGGAAAT GTGCGCGGAA CCCCTATTGT TTTATTTTTC TAAATACATT
4141 CAAATATGTA TCCGCTCATG AGACAATAAC CCTGATAAAT GCTTCAATAA TATTGAAAAA
4201 GGAAGAGTAT GAGTATTCAA CATTTCCGTG TCGCCCTTAT TCCCTTTTTT GCGGCATTTT
4261 GCCTTCCTGT TTTTGCTCAC CCAGAAACGC TGGTGAAAGT AAAAGATGCT GAAGATCAGT
4321 TGGGTGCACG AGTGGGTTAC ATCGAACTGG ATCTCAACAG CGGTAAGATC CTTGAGAGTT
4381 TTCGCCCCGA AGAACGTTTT CCAATGATGA GCACTTTTAA AGTTCTGCTA TGTGGCGCGG
4441 TATTATCCCG TATTGACGCC GGGCAAGAGC AACTCGGTCG CCGCATACAC TATTCTCAGA
4501 ATGACTTGGT TGAGTACTCA CCAGTCACAG AAAAGCATCT TACGGATGGC ATGACAGTAA
4561 GAGAATTATG CAGTGCTGCC ATAACCATGA GTGATAACAC TCGGCGCAAC TTACTTCTGA
4621 CAACGATCGG AGGACCGAAG GAGCTAACCG CTTTTTTGCA CAACATGGGG GATCATGTAA
4681 CTCGCCCTGA TCGTTGGGAA CCGGAGCTGA ATGAAGCCAT ACCAAACGAC GAGCGTGACA
4741 CCACGATGCC TGTAGCAATG GCAACAACGT TCGCGAAACT ATTAACCTGGC GAACTACTTA
4801 CTCTAGCTTC CCGCAACAA TTAATAGACT GGATGGAGGC GGATAAAGTT GCAGGACCAC
4861 TTCTGCGCTC GGCCCTTCCG GCTGGCTGGT TTATTGCTGA TAAATCTGGA GCCGGTGAGC
4921 GTGGGTCTCG CGGTATCATT GCAGCACTGG GGCCAGATGG TAAGCCCTCC CGTATCGTAG
4981 TTATCTACAC GACGGGGAGT CAGGCAACTA TGGATGAACG AAATAGACAG ATCGCTGAGA
5041 TAGGTGCCTC ACTGATTAAG CATTGGTAAC TGTCAGACCA AGTTTACTCA TATATACTTT
5101 AGATTGATTT AAAACTTCAT TTTTAATTTA AAAGGATCTA GGTGAAGATC CTTTTTGATA
5161 ATCTCATGAC CAAAATCCCT TAACGTGAGT TTTTCGTTCCA CTGAGCGTCA GACCCCGTAG
5221 AAAAGATCAA AGGATCTTCT TGAGATCCTT TTTTCTGCG CGTAATCTGC TGCTTGCAAA
5281 CAAAAAACC ACCGCTACCA GCGGTGGTTT GTTTGCCGGA TCAAGAGCTA CCAACTCTTT
5341 TTCCGAAGGT AACTGGCTTC AGCAGAGCGC AGATAACAAA TACTGTCCTT CTAGTGTAGC
5401 CGTAGTTAGG CCACCACTTC AAGAACTCTG TAGCACCGCC TACATACCTC GCTCTGCTAA
5461 TCCTGTTACC AGTGGCTGCT GCCAGTGGCG ATAAGTCGTG TCTTACCGGG TTGGACTCAA
5521 GACGATAGTT ACCGGATAAG GCGCAGCGGT CCGGCTGAAC GGGGGGTTTCG TGCACACAGC
5581 CCAGCTTGGA GCGAACGACC TACACCGAAC TGAGATACCT ACAGCGTGAG CATTGAGAAA
5641 GCGCCACGCT TCCCGAAGGG AGAAAGGCGG ACAGGTATCC GGTAAGCGGC AGGGTCGGAA
5701 CAGGAGAGCG CACGAGGGAG CTTCCAGGGG GAAACGCCTG GTATCTTTAT AGTCTGTGCG
5761 GGTTTCGCCA CCTCTGACTT GAGCGTCGAT TTTTGATGAT CTCGTCAGGG GGGCGGAGCC
5821 TATGAAAAA CGCCAGCAAC GCGGCCTTTT TACGGTTCCT GGCCTTTTGC TGGCCTTTTG
5881 CTCACATGTT CTTTCTGCG TTATCCCTTG ATTCTGTGGA TAACCGTATT ACCGCTTTTG
5941 AGTGAGCTGA TACCGCTCGC CGCAGCCGAA CGACCGAGCG CAGCGAGTCA GTGAGCGAGG
6001 AAGCGGAAGA GCGCCTGATG CGGTATTTTC TCCTTACGCA TCTGTGCGGT ATTTACACACC
6061 GCAGACCAGC CGCGTAACCT GGCAAAATCG GTTACGGTTG AGTAATAAAT GGATGCCCTG
6121 CGTAAGCGGG TGTGGCGCGA CAATAAAGTC TTAACCTGAA CAAAATAGAT CTAACCTATG-

110/240

6181 ACAATAAAGT CTTAAACTAG ACAGAATAGT TGTAAGTGA AATCAGTCCA GTTATGCTGT
6241 GAAAAAGCAT ACTGGACTTT TGTTATGGCT AAAGCAAACCT CTTCATTTTC TGAAGTGCAA
6301 ATTGCCCCGTC GTATTAAAGA GGGGCGTGGC CAAGGGCATG GTAAAGACTA TATTCGCGGC
6361 GTTGTGACAA TTTACCGAAC AACTCCGCGG CCGGGAAGCC GATCTCGGCT TGAACGAATT
6421 GTTAGGTGGC GGTACTTGGG TCGATATCAA AGTGCAATCAC TTCTTCCCGT ATGCCCAACT
6481 TTGTATAGAG AGCCACTGCG GGATCGTCAC CGTAATCTGC TTGCACGTAG ATCACATAAG
6541 CACCAAGCGC GTTGGCCTCA TGCTTGAGGA GATTGATGAG CGCGGTGGCA ATGCCCTGCC
6601 TCCGGTGCTC GCCGGAGACT GCGAGATCAT AGATATAGAT CTCACTACGC GGCTGCTCAA
6661 ACCTGGGCAG AACGTAAGCC GCGAGAGCGC CAACAACCGC TTCTTGCTCG AAGGCAGCAA
6721 GCGCGATGAA TGTCTTACTA CGGAGCAAGT TCCCGAGGTA ATCGGAGTCC GGCTGATGTT
6781 GGGAGTAGGT GGCTACGTCT CCGAACTCAC GACCGAAAAG ATCAAGAGCA GCCCGCATGG
6841 ATTTGACTTG GTCAGGGCCG AGCCTACATG TGCGAATGAT GCCCATACTT GAGCCACCTA
6901 ACTTTGTTTT AGGGCGACTG CCCTGCTGCG TAACATCGTT GCTGCTGCGT AACATCGTTG
6961 CTGCTCCATA ACATCAAACA TCGACCCACG GCGTAACGCG CTTGCTGCTT GGATGCCCGA
7021 GGCATAGACT GTACAAAAA ACAGTCATAA CAAGCCATGA AAACCG

FIGURE 40D

Figure 41A:

pDEST21

**2-Hybrid Vector with
DNA-Binding Domain**

ADH Promoter

700 ~~ttg gcg ctt tgc tat caa gta taa ata gac ctg caa tta tta atc ttt cgt~~
~~aac ggc gaa acg ata gtt cat att tat ctg gac gtt aat aat tag aaa aca~~

751 ~~ttc ctc gtc att gtt ctc gtt ccc ttt ctt cct tgt ttc ttt ttc tgc aca~~
~~aaq gag caa taa caa gaa caa gaa aaa gaa gaa aca aca aca aca aca tct~~

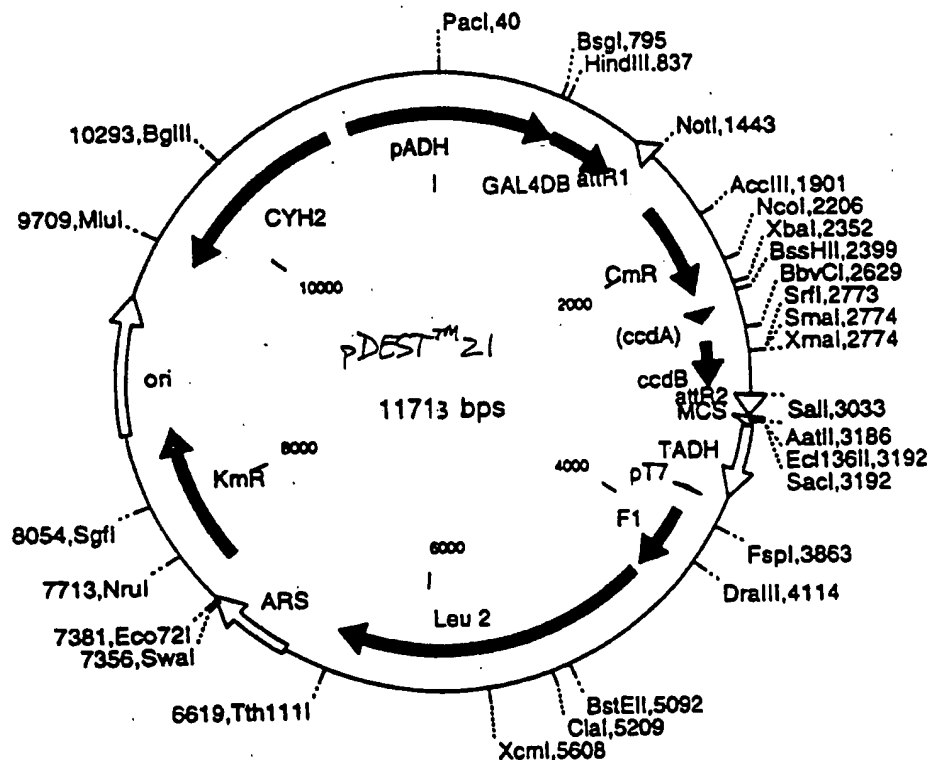
802 ~~ata ttt caa gct ata cca agc ata caa tca act~~ cca agc ttg aag caa gcc
~~tat aaa gtt cga tat ggt tgc tat gtt aat tga~~ ggt tgc aac ttc gtt cgg

853 ~~Start Transla~~ M K L L S S Gal4-DB
~~tcc tga aag atg aag cta ctg tct tct atc gaa caa gca tgc gat att tgc~~
~~agg act ttc tac ttc gat gac aga aga tag ctt gtt cgt acg cta taa acg~~

1261 gaa gag agt agt aac aaa ggt caa aga cag ttg act gta tgc tgc agg tgc
ctt ctc tca tca ttg ttt cca gtt tct gtc aac tga cat agc agc tcc agc

1312 ~~N Q T S L Y K K A~~ att RI
~~aat caa aca agt ttg tac aaa aaa gct gaa cga gaa acg taa aat gat ata~~
~~tta gtt tgt tca aac atg ttt ttt cga ctt gct ctt tgc att tta cta tat~~

Init



112/240

pDEST21 11713 bp (rotated to position 11000)

<u>Location (Base Nos.)</u>		<u>Gene Encoded</u>
857..1322		GAL4DB
1456..1332		attR1
1706..2365		CmR
2485..2569		inactivated ccdA
2707..3012		ccdB
3053..3177		attR2
3716..3735		pT7 (T7 promoter)
3899..4354		f1 (f1 intergenic region)
4414..6642		Leu2
7541..8515		kanR
9668..10958		CYH2
11118..848		pADH (ADH promoter)

1	TTTATTATGT	TACAATATGG	AAGGGAACCT	TACACTTCTC	CTATGCACAT	ATATTAATTA
61	AAGTCCAATG	CTAGTAGAGA	AGGGGGGTAA	CACCCCTCCG	CGCTCTTTTC	CGATTTTTTT
121	CTAAACCGTG	GAATATTTCG	GATATCCTTT	TGTTGTTTCC	GGGTGTACAA	TATGGACTTC
181	CTCTTTTCTG	GCAACCAAAC	CCATACATCG	GGATTCCCTAT	AATACCTTCG	TTGGTCTCCC
241	TAACATGTAG	GTGGCGGAGG	GGAGATATAC	AATAGAACAG	ATACCAGACA	AGACATAATG
301	GGCTAAACAA	GACTACACCA	ATTACACTGC	CTCATTGATG	GTGGTACATA	ACGAACATAAT
361	ACTGTAGCCC	TAGACTTGAT	AGCCATCATC	ATATCGAAGT	TTCACTACCC	TTTTTCCATT
421	TGCCATCTAT	TGAAGTAATA	ATAGGCGCAT	GCAACTTCTT	TTCTTTTTTT	TTCTTTTCTC
481	TCTCCCCCGT	TGTTGTCTCA	CCATATCCGC	AATGACAAAA	AAAATGATGG	AAGACACTAA
541	AGGAAAAAAT	TAACGACAAA	GACAGCACCA	ACAGATGTCG	TTGTTCCAGA	GCTGATGAGG
601	GGTATCTTCG	AACACACGAA	ACTTTTTCCT	TCCTTCATTC	ACGCACACTA	CTCTCTAATG
661	AGCAACGGTA	TACGGCCTTC	CTTCCAGTTA	CTTGAATTTG	AAATAAAAAA	AGTTTGCCGC
721	TTTGCTATCA	AGTATAAATA	GACCTGCAAT	TATTAATCTT	TTGTTTCCTC	GTCATTGTTC
781	TCGTTCCCTT	TCTTCCTTGT	TTCTTTTCT	GCACAATATT	TCAAGCTATA	CCAAGCATAC
841	AATCAACTCC	AAGCTTGAAG	CAAGCCTCCT	GAAAGATGAA	GCTACTGTCT	TCTATCGAAC
901	AAGCATGCGA	TATTTGCCGA	CTTAAAAAGC	TCAAGTGCTC	CAAAGAAAAA	CCGAAGTGCG
961	CCAAGTGTCT	GAAGAACAAC	TGGGAGTGTC	GCTACTCTCC	CAAACCAAAA	AGGTCTCCGC
1021	TGACTAGGGC	ACATCTGACA	GAAGTGGAAT	CAAGGCTAGA	AAGACTGGAA	CAGCTATTTT
1081	TACTGATTTT	TCCTCGAGAA	GACCTTGACA	TGATTTTGAA	AATGGATTCT	TTACAGGATA
1141	TAAAAGCATT	GTAAACAGGA	TTATTTGTAC	AAGATAATGT	GAATAAAGAT	GCCGTCACAG
1201	ATAGATTGGC	TTCAAGTGAG	ACTGATATGC	CTCTAACATT	GAGACAGCAT	AGAATAAGTG
1261	CGACATCATC	ATCGGAAGAG	AGTAGTAACA	AAGGTCAAAG	ACAGTTGACT	GTATCTCGCA
1321	GGTCGAATCA	AACAAGTTTG	TACAAAAAAG	CTGAACGAGA	AACGTAAAAT	GATATAAATA
1381	TCAATATATT	AAATTAGATT	TTGCATAAAA	AACAGACTAC	ATAATACTGT	AAAACACAAC
1441	ATATCCAGTC	ACTATGGCGG	CCGCTAAGTT	GGCAGCATCA	CCCGACGCAC	TTTGCGCCGA
1501	ATAAATACCT	GTGACGGAAG	ATCACTTCGC	AGAATAAATA	AATCCTGGTG	TCCCTGTTGA
1561	TACCGGGAAG	CCCTGGGCCA	ACTTTTGGCG	AAAATGAGAC	GTTGATCGGC	ACGTAAGAGG
1621	TTCCAACCTT	CACCATAATG	AAATAAGATC	ACTACCGGGC	GTATTTTTTG	AGTTATCGAG
1681	ATTTTCAGGA	GCTAAGGAAG	CTAAAAATGA	GAAAAAAATC	ACTGGATATA	CCACCGTTGA
1741	TATATCCCAA	TGGCATCGTA	AAGAACATTT	TGAGGCATTT	CAGTCAGTTG	CTCAATGTAC
1801	CTATAACCAG	ACCGTTCAGC	TGGATATTAC	GGCCTTTTTA	AAGACCGTAA	AGAAAAATAA
1861	GCACAAGTTT	TATCCGGCCT	TTATTCACAT	TCTTGCCCGC	CTGATGAATG	CTCATCCGGA
1921	ATTCCGTATG	GCAATGAAAG	ACGGTGAGCT	GGTGATATGG	GATAGTGTTC	ACCCTTGTTA
1981	CACCGTTTTT	CATGAGCAAA	CTGAAACGTT	TTCATCGCTC	TGGAGTGAAT	ACCACGACGA
2041	TTTCCGGCAG	TTTCTACACA	TATATTCGCA	AGATGTGGCG	TGTTACGGTG	AAAACCTGGC
2101	CTTTTCCCT	AAAGGGTTTA	TTGAGAATAT	GTTTTTCGTC	TCAGCCAATC	CCTGGGTGAG
2161	TTTCACCACT	TTTGATTAA	ACGTGGCCAA	TATGGACAAC	TTCTTCGCCC	CCGTTTTTAC
2221	CATGGGCAAA	TATTATACGC	AAGGCGACAA	GGTGCTGATG	CCGCTGGCGA	TTCAGGTTCA
2281	TCATGCCGTC	TGTGATGGCT	TCCATGTCGG	CAGAATGCTT	AATGAATTAC	AACAGTACTG
2341	CGATGAGTGG	CAGGGCGGGG	CGTAATCTAG	AGGATCCGGC	TTACTAAAAG	CCAGATAACA
2401	GTATGCGTAT	TTGCGCGCTG	ATTTTTGCGG	TATAAGAATA	TATACTGATA	TGTATACCGG-

FIGURE 413

2461 AAGTATGTCA AAAAGAGGTG TGCTATGAAG CAGCGTATTA CAGTGACAGT TGACAGCGAC
2521 AGCTATCAGT TGCTCAAGGC ATATATGATG TCAATATCTC CGGTCTGGTA AGCACAACCA
2581 TGCAGAATGA AGCCCGTCGT CTGCGTGCCG AACGCTGGAA AGCGGAAAAT CAGGAAGGGA
2641 TGGCTGAGGT CGCCCGGTTT ATTGAAATGA ACGGCTCTTT TGCTGACGAG AACAGGGACT
2701 GGTGAAATGC AGTTTAAGGT TTACACCTAT AAAAGAGAGA GCCGTTATCG TCTGTTTGTG
2761 GATGTACAGA GTGATATTAT TGACACGCCC GGGCGACGGA TGGTGATCCC CCTGGCCAGT
2821 GCACGTCTGC TGTCAGATAA AGTCTCCCGT GAACTTTACC CGGTGGTGCA TATCGGGGAT
2881 GAAAGCTGGC GCATGATGAC CACCGATATG GCCAGTGTGC CGGTCTCCGT TATCGGGGAA
2941 GAAGTGGCTG ATCTCAGCCA CCGCGAAAAT GACATCAAAA ACGCCATTAA CCTGATGTTT
3001 TGGGGAATAT AAATGTCAGG CTCCCTTATA CACAGCCAGT CTGCAGGTCG ACCATAGTGA
3061 CTGGATATGT TGTGTTTTAC AGTATTATGT AGTCTGTTTT TTATGCAAAA TCTAATTTAA
3121 TATATTGATA TTTATATCAT TTTACGTTTC TCGTTCAGCT TTCTTGTAACA AAGTGGTTTG
3181 ATGGCCGCTA AGTAAGTAAG ACGTCGAGCT CTAAGTAAGT AACGGCCGCC ACCGCGGTGG
3241 AGCTTTGGAC TTCTTCGCCA GAGGTTTGGT CAAGTCTCCA ATCAAGGTTG TCGGCTTGTC
3301 TACCTTGCCA GAAATTTACG AAAAGATGGA AAAGGGTCAA ATCGTTGGTA GATACGTTGT
3361 TGACACTTCT AAATAAGCGA ATTTCTTATG ATTTATGATT TTTATTATTA AATAAGTTAT
3421 AAAAAAATA AGTGATACA AATTTTAAAG TGACTCTTAG GTTTTAAAC GAAAATTCTT
3481 ATTCTTGAGT AACTCTTTCC TGTAGGTCAG GTTGCTTTCT CAGGTATAGC ATGAGGTCGC
3541 TCTTATTGAC CACACCTCTA CCGGCATGCC GAGCAAATGC CTGCAAATCG CTCCCCATT
3601 CACCAATTG TAGATATGCT AACTCCAGCA ATGAGTTGAT GAATCTCGGT GTGTATTTTA
3661 TGCTCTCAGA GGACAATACC TGTTGTAATC GTTCTTCCAC ACGGATCCCA ATTGCGCCTA
3721 TAGTGAGTCG TATTACAATT CACTGGCCGT CGTTTTACAA CGTCGTGACT GGGAAAACCC
3781 TGGCGTTACC CAACTTAATC GCCTTGACAG ACATCCCCCT TTCGCCAGCT GCGGTAATAG
3841 CGAAGAGGCC CGCACCGATC GCCCTTCCCA ACAGTTGCGC AGCCTGAATG GCGAATGGAC
3901 GCGCCCTGTA GCGGCGCATT AAGCGCGGCG GGTGTGGTGG TTACGCGCAG CGTGACCGCT
3961 ACACTTGCCA GCGCCCTAGC GCCCGCTCCT TTCGCTTTCT TCCCTTCCTT TCTCGCCACG
4021 TTCGCGGCTT TCCCCGTC AAGCTCTAAT CCGGGGCTCC CTTAGGGTT CCGATTTAGT
4081 GCTTTACGGC ACCTCGACCC CAAAAAATT GATTAGGGTG ATGGTTCACG TAGTGGGCCA
4141 TCGCCCTGAT AGACGGTTTT TCGCCCTTTG ACGTTGGAGT CCACGTTCTT TAATAGTGGA
4201 CTCTTGTTCC AAAGTGAAC AACACTCAAC CCTATCTCGG TCTATTCTTT TGATTTATAA
4261 GGGATTTTGC CGATTTCCGC CTATTGGTTA AAAAATGAGC TGATTTAACA AAAATTTAAC
4321 GCGAATTTTA ACAAATATT AACGTTTACA ATTTCTGAT GCGGTATTTT CTCCTTACGC
4381 ATCTGTGCGG TATTTACAC CGCATATCGA CCGGTGAGG AGAATTCTA GTATATCCAC
4441 ATACCTAATA TTATTGCCTT ATTAAAAATG GAATCGGAAC AATTACATCA AAATCCACAT
4501 TCTCTTCAAA ATCAATTGTC CTGTACTTCC TTGTTTCATG GTGTTCAAAA ACGTTATATT
4561 TATAGGATAA TTATACTCTA TTTCTCAACA AGTAATTGGT TGTTTGGCCG AGCGGTCTAA
4621 GCGCGCTGAT TCAAGAAATA TCTTGACCGC AGTTAACTGT GGGAAATACT AGGTATCGTA
4681 AGATGCAAGA GTTCGAATCT CTTAGCAACC ATTATTTTTT TCCTCAACAT AACGAGAACA
4741 CACAGGGGCG CTATCGCACA GAATCAAATT CGATGACTGG AAATTTTTTG TTAATTTTCA
4801 AGGTCGCGCT ACGCATATAC CTTTTTCAAC TGAAAAATG GGAGAAAAAG GAAAGGTGAG
4861 AGGCCGGAAC CGGCTTTTCA TATAGAATAG AGAAGCGTTC ATGACTAAAT GCTTGCATCA
4921 CAATACTTGA AGTTGACAAT ATTATTTAAG GACCTATTGT TTTTCCAAT AGGTGGTTAG
4981 CAATCGTCTT ACTTTCTAAC TTTTCTTACC TTTTACATTT CAGCAATATA TATATATATT
5041 TCAAGGATAT ACCATTCTAA TGTCTGCCCC TATGTCTGCC CTTAAGAAGA TCGTCGTTTT
5101 GCCAGGTGAC CACGTTGGTC AAGAAATCAC AGCCGAAGCC ATTAAGGTTT TTAAGCTAT
5161 TTCTGATGTT CGTTCCAATG TCAAGTTCGA TTTGCGAAAAT CATTTAATTG GTGGTGCTGC
5221 TATCGATGCT ACAGGTGTCC CACTTCCAGA TGAGGCGCTG GAAGCCTCCA AGAAGGTTGA
5281 TGCCGTTTTG TTAGGTGCTG TGGGTGGTCC TAAATGGGGT ACCGCTAGTG TTAGACCTGA
5341 ACAAGGTTTA CTAAAAATCC GTAAAGAACT TCAATTGTAC GCCAACTTAA GACCATGTAA
5401 CTTTGTCATCC GACTCTCTTT TAGACTTATC TCCAATCAAG CCACAATTTG CTAAAGGTAC
5461 TGACTTCGTT GTTGTGAGAG AATTAGTGGG AGGTATTTAC TTTGGTAAGA GAAAGGAAGA
5521 CGATGGTGAT GGTGTGCGTT GGGATAGTGA ACAATACACC GTTCCAGAAG TGCAAGAAT
5581 CACAAGAATG GCCGCTTTCA TGGCCCTACA ACATGAGCCA CCATTGCCTA TTTGGTCTTT
5641 GGATAAAGCT AATGTTTTGG CCTCTTCAAG ATTATGGAGA AAAACTGTGG AGGAAACCAT
5701 CAAGAACGAA TTCCCTACAT TGAAGGTTCA ACATCAATTG ATTGATCTG CCGCATGAT
5761 CCTAGTTAAG AACCCAACCC ACCTAAATGG TATTATAATC ACCAGCAACA GTTTGGTGA
5821 TATCATCTCC GATGAAGCCT CCGTTATCCC AGGTTCCTTG GGTGTTGTGC CATCTGCGTC
5881 CTTGGCCTCT TTGCCAGACA AGAACACCGC ATTTGGTTTG TACGAACCAT GCCACGGTTC

Figure 41C

5941 TGCTCCAGAT TTGCCAAAGA ATAAGGTTGA CCCTATCGCC ACTATCTTGT CTGCTGCAAT
6001 GATGTTGAAA TTGTCATTGA ACTTGCCTGA AGAAGGTAAG GCCATTGAAG ATGCAGTTAA
6061 AAAGGTTTTG GATGCAGGTA TCAGAACTGG TGATTTAGGT GGTTCACA GTACCACCGA
6121 AGTCGGTGAT GCTGTGCGCG AAGAAGTTAA GAAAATCCTT GCTTAAAAAG ATTCTCTTTT
6181 TTTATGATAT TTGTACATAA ACTTTATAAA TGAAATTCAT AATAGAAACG ACACGAAATT
6241 ACAAAATGGA ATATGTTTAT AGGGTAGACG AAACATATATA CGCAATCTAC ATACATTTAT
6301 CAAGAAGGAG AAAAAGGAGG ATAGTAAAGG AATACAGGTA AGCAAATTGA TACTAATGGC
6361 TCAACGTGAT AAGGAAAAAG AATTGCACTT TAACATTAAT ATTGACAAGG AGGAGGGCAC
6421 CACACAAAAA GTTAGGTGTA ACAGAAAATC ATGAAACTAC GATTCTAAT TTGATATTGG
6481 AGGATTTTCT CTAACAAAAA AAAAATACAA CAAATAAAAA AACTCAATG ACCTGACCAT
6541 TTGATGGAGT TTAAGTCAAT ACCTTCTTGA ACCATTTCCC ATAATGGTGA AAGTTCCCTC
6601 AAGAATTTTA CTCTGTGAGA AACGGCCTTA CGACGTAGTC GATATGGTGC ACTCTCAGTA
6661 CAATCTGCTC TGATGCCGCA TAGTTAAGCC AGCCCCGACA CCCGCCAACA CCCGCTGACG
6721 CGCCCTGACG GGCTTGTCTG CTCCCGCAT CCGCTTACAG ACAAGCTGTG ACCGTCTCCG
6781 GGAGCTGCAT GTGTGAGAGG TTTTCACCGT CATCACCGAA ACGCGCGAGA AGGAGGGCAC
6841 TCGTGATACG CCTATTTTAA TAGGTAAATG TCATGATAAT AATGGTTTCT TAGGACGGAT
6901 CGCTTGCTG TAACTTACAC GCGCTCGTA TCTTTTAAATG ATGGAATAAT TTGGGAATTT
6961 ACTCTGTGTT TATTTATTTT TATGTTTGT ATTTGGATT TAGAAAGTAA ATAAAGAAGG
7021 TAGAAGAGTT ACGGAATGAA GAAAAAATAA TAAACAAAGG TTTAAAAAAT TTCAACAAAA
7081 AGCGTACTTT ACATATATAT TTATTAGACA AGAAAAGCAG ATTAAATAGA TATACATTCTG
7141 ATTAACGATA AGTAAATGT AAAATCACAG GATTTTCGTG TGTGGTCTTC TACACAGACA
7201 AGATGAAACA ATTCGGCATT AATACCTGAG AGCAGGAAGA GCAAGATAAA AGGTAGTATT
7261 TGTTGGCGAT CCCCTAGAG TCTTTTACAT CTTCCGAAAA CAAAAACTAT TTTTCTTTA
7321 ATTTCTTTTT TTACTTTCTA TTTTAAATTT ATATATTTAT ATTAAAAAAT TTAAATTATA
7381 ATTATTTTTA TAGCACGTGA TGAAAAGGAC CCAGGTGGCA CTTTTCGGGG AAATGTGCGC
7441 GGAACCCCTA TTTGTTTATT TTTCTAAATA CATCAATA TGTATCCGCT CATGAGACAA
7501 TAACCCTGAT AAATGCTTCA ATAATCTGCA GCTCTGGCCC GTGTCTCAA ATCTCTGATG
7561 TTACATGCA CAAGATAAAA ATATATCAT ATGAACAATA AAAGTGTCTG CTTACATAAA
7621 CAGTAATACA AGGGGTGTTA TGAGCCATAT TCAACGGGAA ACGTCTTGCT GGAGGCCGCG
7681 ATTAAATTCC AACATGGATG CTGATTTATA TGGGTATAAA TGGGCTCGCG ATAATGTCGG
7741 GCAATCAGGT GCGACAATCT TTCGATTGTA TGGGAAGCCC GATGCGCCAG AGTTGTTTCT
7801 GAAACATGGC AAAGGTAGCG TTGCCAATGA TGTTACAGAT GAGATGGTCA GACTAACTG
7861 GCTGACGGAA TTTATGCCTC TTCCGACCAT CAAGCATTTT ATCCGTACTC CTGATGATGC
7921 ATGGTTACTC ACCACTGCGA TCCGCGGGAA AACAGCATT CAGGTATTAG AAGAATATCC
7981 TGATTGAGT GAAAATATTG TTGATGCGCT GGCAGTGTTC CTGCGCCGCT TGCATTCGAT
8041 TCCTGTTTGT AATTGTCCTT TTAACAGCGA TCGCGTATTT CGTCTCGCTC AGCGCGATC
8101 ACGAATGAAT AACGGTTTGG TTGATGCGAG TGATTTTGAT GACGAGCGTA ATGGCTGGCC
8161 TGTTGAACAA GTCTGGAAAG AAATGCATAC GCTTTTGCCA TTCTCACCGG ATTCAGTCGT
8221 CACTCATGGT GATTTCTCAC TTGATAACCT TATTTTGGAC GAGGGGAAAT TAATAGGTTG
8281 TATTGATGTT GGACGAGTCG GAATCGCAGA CCGATACCAG GATCTTGCCA TCCTATGGAA
8341 CTGCTCGGT GAGTTTCTC CTTCATTACA GAAACGGCTT TTTCAAAAAT ATGGTATTGA
8401 TAATCCTGAT ATGAATAAAT TGCAGTTTCA TTTGATGCTC GATGAGTTT TCTAATCAGA
8461 ATTGGTTAAT TGGTTGTAAC ACTGGCAGAG CATTACGCTG ACTTGACGGG ACGGCGCATG
8521 ACCAAAATCC CTTAACGTGA GTTTTCGTTT CACTGAGCGT CAGACCCCGT AGAAAAGATC
8581 AAAGGATCTT CTTGAGATCC TTTTCTCTG CGCGTAATCT GCTGCTTGCA AACAAAAAAA
8641 CCACCGCTAC CAGCGGTGGT TTGTTTGCCG GATCAAGAGC TACCAACTCT TTTCCGAAG
8701 GTAACCTGGT TCAGCAGAGC GCAGATACCA AATACTGTCC TTCTAGTGTA GCCGTAGTTA
8761 GGCCACCACT TCAAGAACTC TGTAAGCACC CCTACATACC TCGCTCTGCT AATCCTGTTA
8821 CCAGTGGCTG CTGCCAGTGG CGATAAGTCA TGTCTTACCG GGTGGACTC AAGACGATAG
8881 TTACCGGATA AGGCGCAGCG GTCGGGCTGA ACGGGGGGTT CGTGCACACA GCCCAGCTTG
8941 GAGCGAACGA CCTACACCGA ACTGAGATAC CTACAGCGTG AGCATTGAGA AAGCGCCACG
9001 CTTCCCGAAG GGAGAAAGGC GGACAGGTAT CCGGTAAGCG GCAGGTCGG AACAGGAGAG
9061 CGCACGAGGG AGCTTCCAGG GGGGAACGCC TGGTATCTTT ATAGTCTGTG CGGGTTTCGC
9121 CACCTCTGAC TTGAGCGTCG ATTTTGTGTA TGCTCGTCAG GGGGGCCGAG CCTATGGAAA
9181 AACGCCAGCA ACGCGGCTT TTTACGGTTC CTGGCCTTTT GCTGGCCTTT TGCTCACATG
9241 TTCTTTCTG CGTTATCCCC TGATTCTGTG GATAACCGTA TTACCGCTT TGAGTGAGCT
9301 GATACCGCTC GCCGACCGC AACGACCGAG CGCAGCGAGT CAGTGAGCGA GGAAGCGGAA
9361 GAGCGCCCAA TACGCAAAACC GCCTCTCCCC GCGCGTTGGC CGATTCATTA ATGCAGCTGG-

FIGURE 4D

115/240

9421 CACGACAGGT TTCCCGACTG GAAAGCGGGC AGTGAGCGCA ACGCAATTAA TGTGAGTTAC
9481 CTCACCTCATT AGGCACCCCA GGCTTTACAC TTTATGCTTC CGGCTCCTAT GTTGTGTGGA
9541 ATTGTGAGCG GATAACAATT TCACACAGGA AACAGCTATG ACCATGATTA CGCCAAGCTC
9601 GGAATTAACC CTCACTAAAG GGAACAAAAG CTGGTACCGA TCCCGAGCTT TGCAAATTAA
9661 AGCCTTCGAG CGTCCCAAAA CCTTCTCAAG CAAGGTTTTT AGTATAATGT TACATGCGTA
9721 CACGCGTCTG TACAGAAAAA AAAGAAAAAT TTGAAATATA AATAACGTTT TTAATACTAA
9781 CATAACTATA AAAAAATAAA TAGGGACCTA GACTTCAGGT TGTCTAACTC CTTCTTTTTC
9841 GGTTAGAGCG GATGTGGGGG GAGGGCGTGA ATGTAAGCGT GACATAACTA ATTACATGAT
9901 ATCGACAAAG GAAAAGGGGC CTGTTTACTC ACAGGCTTTT TTCAAGTAGG TAATTAAGTC
9961 GTTCTGTCT TTTTCTTCT TCAACCCACC AAAGGCCATC TTGGTACTTT TTTTTTTTTT
10021 TTTTTTTTTT TTTTTTTTTT TTTTTTTTTT TTTTTTTTTT TTTTTTTTTT TTTTTTTTTT
10081 TTTTTTTTTT TTTTTTTTTT TCATAGAAAT AATACAGAAG TAGATGTTGA ATTAGATTAA
10141 ACTGAAGATA TATAATTTAT TGGAAAATAC ATAGAGCTTT TTGTTGATGC GCTTAAGCGA
10201 TCAATTCAAC AACACCACCA GCAGCTCTGA TTTTCTTCTC AGCCAACCTG GAGACGAATC
10261 TAGCTTTGAC GATAACTGGA ACATTTGGAA TTCTACCTTT ACCCAAGATC TTACCGTAAC
10321 CGGCTGCCAA AGTGTCAATA ACTGGAGCAG TTTCTTAGA AGCAGATTTC AAGTATTGGT
10381 CTCTCTGTG TCTGCGGATC AATGTCCACA ATTTGTCCAA GTTCAAGACT GGCTTCCAGA
10441 AATGAGCTTG TTGCTTGTGG AAGTATCTCA TACCAACCTT ACCGAAATAA CCTGGATGGT
10501 ATTTATCCAT GTTAATTCTG TGGTGATGTT GACCACCGGC CATACCTCTA CCACCGGGGT
10561 GCTTCTGTG CTTACCGATA CGACCTTTAC CGGCTGAGAC GTGACCTCTG TGCTTTCTAG
10621 TCTTAGTGAA TCTGGAAGGC ATTCTTGATT AGTTGGATGA TTGTTCTGGG ATTTAATGCA
10681 AAAATCACTT AAGAAGGAAA ATCAACGGAG AAAGCAAACG CCATCTTAAA TATACGGGAT
10741 ACAGATGAAA GGGTTTGAAC CTATCTGGAA AATAGCATT AACAAGCGAA AAAGTGGAG
10801 GAAAATTGTT TGGCTCTCTG CGGGCTATT CCGCGCCAGA GGAAAATAGG AAAAATAACA
10861 GGGCATTAGA AAAATAATTT TGATTTTGGT AATGTGTGGG TCCTGGTGTA CAGATGTTAC
10921 ATTGTTTACA GTACTCTTGT TTTTGCTGTG TTTTTCGATG AATCTCCAAA ATGGTTGTTA
10981 GCACATGGAA GAGTCACCGA TGCTAAGTTA TCTCTATGTA AGCTACGTGG CGTGACTTTT
11041 GATGAAGCCG CACAAGAGAT ACAGGATTGG CAACTGCAAA TAGAATCTGG GGATCCCCC
11101 TCGAGATCCG GGATCGAAGA AATGATGGTA AATGAAATAG GAAATCAAGG AGCATGAAGG
11161 CAAAAGACAA ATATAAGGGT CGAACGAAAA ATAAAGTGAA AAGTGTGAT ATGATGTATT
11221 TGGCTTTGCG GCGCCGAAAA AACGAGTTTA CGCAATTGCA CAATCATGCT GACTCTGTGG
11281 CGGACCCGCG CTCTTGCCGG CCCGGCGATA ACGCTGGGCG TGAGGCTGTG CCCGGCGGAG
11341 TTTTTTGC GC CTGCATTTTT CAAGGTTTAC CCTGCGCTAA GGGGCGAGAT TGGAGAAGCA
11401 ATAAGATGC CGGTTGGGGT TGCGATGATG ACGACCACGA CAACTGGTGT CATTATTTAA
11461 GTTGCCGAAA GAACCTGAGT GCATTTGCAA CATGAGTATA CTAGAAGAAT GAGCCAAGAC
11521 TTGCGAGACG CGAGTTTGCC GGTGGTGCGA ACAATAGAGC GACCATGACC TTGAAGGTGA
11581 GACGCGCATA ACCGCTAGAG TACTTTGAAG AGGAAACAGC AATAGGGTTG CTACCAGTAT
11641 AAATAGACAG GTACATACAA CACTGGAAAT GGTTGTCTGT TTGAGTACGC TTTCAATTCA
11701 TTTGGGTGTG CAC

FIGURE 415

Figure 42A:

PDST22

2-Hybrid Vector with
Activation Domain

657 acg cac act act ctc taa tga gca acg gta tac ggc ctt cct tcc agt tac
tgc gtg tga tga gag att act cgt tgc cat atg ccg gaa gga agg tca atg

708 ttg aat ttg aaa taa aaa aag ttt gcc gct ttg cta tca agt ata aat aga
aac tta aac ttt att ttt ttc aaa cgg cga aac gat agt tca tat tta tct

759 cct gca att att aat ctt ttg ttt cct cgt cat tgt tct cgt tcc ctt tct
gga cgt taa taa tta gaa aac aaa gga gca gta aca aga gca agg gaa aga

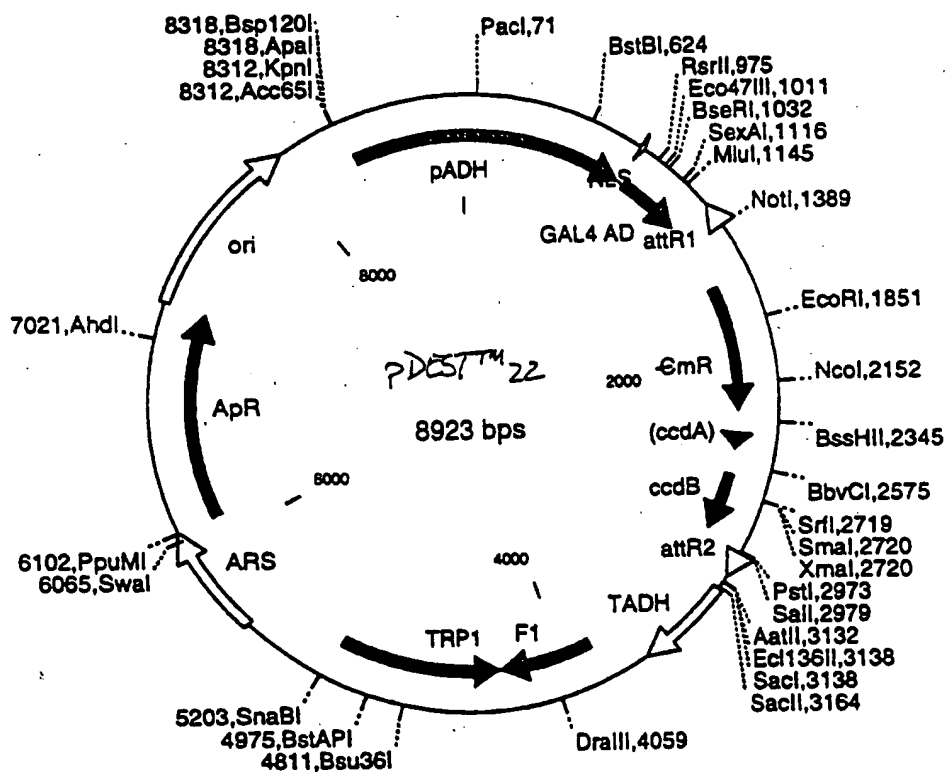
810 ~~ccc/ttg ttt ctt ttt ctt cat aat att tca apc tat acc aag cat aca apc~~
~~agg aac aac gaa aaa gac dtg cta taa agt tca ata tgg ttc gta tct tag~~

861 ~~aac/tcc aag ctt atg ccc aag aag aag cgg aag gtc tgg agc ggc gcc aat~~
~~ctg agg ttc gaa tac ggg ttc ttc ttc gcc ttc cag agc tgg ccg cgg ttg~~

1218 gaa gat acc cca cca aac cca aaa aaa ^{D G G S N Q T S} gag ggt ggg tgg aat caa aca agt
ctt cta tgg ggt ggt ttg ggt ttt ttt ctc cca ccc agc tta gtt tgt tca

1269 ~~ttg tac aaa aaa gct gaa cga gaa acg taa a~~
~~aac atg ttt ttt cga ctt gct ctt tgc att t~~

ADH Promoter
GAL4-AD
Start Translation
IntV



117/240

pDEST22 8923 bp

<u>Location (Base Nos.)</u>		<u>Gene Encoded</u>
904..1248		GAL4 AD
1388..1264		attR1
1638..2297		CmR
2417..2501		inactivated ccdA
2639..2944		ccdB
2985..3109		attR2
3831..4318		f1 (f1 intergenic region)
4334..5176		TRP1
6110..7194		ampR
8344..866		pADH (yeast ADH promoter)

1	TTCATTGGG	TGTGCACTTT	ATTATGTTAC	AATATGGAAG	GGAACTTTAC	ACTTCTCCTA
61	TGCACATATA	TTAATTAAAG	TCCAATGCTA	GTAGAGAAGG	GGGGTAACAC	CCCTCCGCGC
121	TCTTTTCCGA	TTTTTTTCTA	AACCGTGGAA	TATTTCCGGAT	ATCCTTTTGT	TGTTTTCCGGG
181	TGTACAATAT	GGACTTCCTC	TTTTCTGGCA	ACCAAACCCA	TACATCGGGA	TTCTTATAAT
241	ACCTTCGTTG	GTCTCCCTAA	CATGTAGGTG	GCGGAGGGGA	GATATACAAT	AGAACAGATA
301	CCAGACAAGA	CATAATGGGC	TAAACAAGAC	TACACCAATT	ACACTGCCTC	ATTGATGGTG
361	GTACATAACG	AACTAATACT	GTAGCCCTAG	ACTTGATAGC	CATCATCATA	TCGAAGTTTC
421	ACTACCCTTT	TTCCATTTGC	CATCTATTGA	AGTAATAATA	GGCGCATGCA	ACTTCTTTTC
481	TTTTTTTTTC	TTTTCTCTCT	CCCCCGTTGT	TGTCTCACCA	TATCCGCAAT	GACAAAAAAA
541	ATGATGGAAG	ACACTAAAGG	AAAAAATTAA	CGACAAAGAC	AGCACCAACA	GATGTCGTTG
601	TTCCAGAGCT	GATGAGGGGT	ATCTTCGAAC	ACACGAAACT	TTTTCTTCC	TTTATTCCAG
661	CACACTACTC	TCTAATGAGC	AACGGTATAC	GGCCTTCCTT	CCAGTTACTT	GAATTTGAAA
721	TAAAAAAGT	TTGCCGCTTT	GCTATCAAGT	ATAAATAGAC	CTGCAATTAT	TAATCTTTTG
781	TTTCCTCGTC	ATTGTTCTCG	TTCCCTTTCT	TCCTTGTTTC	TTTTTCTGCA	CAATATTTCA
841	AGCTATACCA	AGCATACAAT	CAACTCCAAG	CTTATGCCCA	AGAAGAAGCG	GAAGGTCTCG
901	AGCGGCGCCA	ATTTAATCA	AAGTGGGAAT	ATTGCTGATA	GCTCATTGTC	CTTCACTTTC
961	ACTAACAGTA	GCAACGGTCC	GAACCTCATA	ACAACTCAA	CAAATTCTCA	AGCGCTTTCA
1021	CAACCAATTG	CCTCCTCTAA	CGTTCATGAT	AACTTCATGA	ATAATGAAAT	CACGGCTAGT
1081	AAAATTGATG	ATGGTAATAA	TTCAAAACCA	CTGTACCTG	GTTGGACGGA	CCAAACTGCG
1141	TATAACGCGT	TTGGAATCAC	TACAGGGATG	TTTAATACCA	CTACAATGGA	TGATGTATAT
1201	AACTATCTAT	TCGATGATGA	AGATACCCCA	CCAAACCCAA	AAAAAGAGGG	TGGGTGCAAT
1261	CAACAAGTT	TGTACAAAA	AGCTGAACGA	GAAACGTAAA	ATGATATAAA	TATCAATATA
1321	TTAAATTAGA	TTTTGCATAA	AAAACAGACT	ACATAATACT	GTAAAAACACA	ACATATCCAG
1381	TCACTATGGC	GGCCGCTAAG	TTGGCAGCAT	CACCCGACGC	ACTTTGCGCC	GAATAAATAC
1441	CTGTGACGGA	AGATCACTTC	GCAGAATAAA	TAAATCCTGG	TGTCCCTGTT	GATACCGGGA
1501	AGCCCTGGGC	CAACTTTTGG	CGAAAATGAG	ACGTTGATCG	GCACGTAAGA	GGTTCCAAC
1561	TTCAACATAA	TGAAATAAGA	TCACTACCGG	GCGTATTTTT	TGAGTTATCG	AGATTTTCAG
1621	GAGCTAAGGA	AGCTAAATG	GAGAAAAAAA	TCACTGGATA	TACCACCGTT	GATATATCCC
1681	AATGGCATCG	TAAAGAACAT	TTTGAGGCAT	TTCAGTCAGT	TGCTCAATGT	ACCTATAACC
1741	AGACCGTTCA	GCTGGATATT	ACGGCCTTTT	TAAAGACCGT	AAAGAAAAAT	AAGCACAAAT
1801	TTTATCCGGC	CTTTATTAC	ATTCTTGCCC	GCCTGATGAA	TGCTCATCCG	GAATTCCGTA
1861	TGGCAATGAA	AGACGGTGAG	CTGGTGATAT	GGGATAGTGT	TCACCCTTGT	TACACCGTTT
1921	TCCATGAGCA	AACTGAAACG	TTTTCATCGC	TCTGGAGTGA	ATACCACGAC	GATTTCCGGC
1981	AGTTTCTACA	CATATATTCG	CAAGATGTGG	CGTGTACGG	TGAAAACCTG	GCCTATTTCC
2041	CTAAAGGGTT	TATTGAGAAT	ATGTTTTTCG	TCTCAGCCAA	TCCTTGGGTG	AGTTTCACCA
2101	GTTTTGATTT	AAACGTGGCC	AATATGGACA	ACTTCTTCGC	CCCCGTTTTT	ACCATGGGCA
2161	AATATTATAC	GCAAGGCGAC	AAGGTGCTGA	TGCCGCTGGC	GATTCAAGTT	CATCATCCCG
2221	TCTGTGATGG	CTTCCATGTC	GGCAGAATGC	TTAATGAATT	ACAACAGTAC	TGCGATGAGT
2281	GGCAGGCGCG	GGCGTAATCT	AGAGGATCCG	GCTTACTAAA	AGCCAGATAA	CAGTATGCGT
2341	ATTTGCGCGC	TGATTTTTGC	GGTATAAGAA	TATATACTGA	TATGTATACC	CGAAGTATGT
2401	CAAAAGAGAG	TGTGCTATGA	AGCAGCGTAT	TACAGTGACA	GTTGACAGCG	ACAGCTATCA
2461	GTTGCTCAAG	GCATATATGA	TGTCAATATC	TCCGGTCTGG	TAAGCACAA	CATGCAGAAT
2521	GAAGCCCGTC	GTCTGCGTGC	CGAACGCTGG	AAAGCGGAAA	ATCAGGAAGG	GATGGCTGAG-

FIGURE 425

2581 GTCGCCCGGT TTATTGAAAT GAACGGCTCT TTTGCTGACG AGAACAGGGA CTGGTGAAAT
2641 GCAGTTTAAG GTTTACACCT ATAAAAGAGA GAGCCGTTAT CGTCTGTTT TGGATGTACA
2701 GAGTGATATT ATTGACACGC CCGGGCGACG GATGGTGATC CCCCTGGCCA GTGCACGTCT
2761 GCTGTGAGAT AAAGTCTCCC GTGAACTTTA CCCGGTGGTG CATATCGGG ATGAAAGCTG
2821 GCGCATGATG ACCACCGATA TGGCCAGTGT GCCGGTCTCC GTTATCGGG AAGAAGTGGC
2881 TGATCTCAGC CACCGCGAAA ATGACATCAA AAACGCCATT AACCTGATGT TCTGGGGAAT
2941 ATAAATGTCA GGCTCCCTTA TACACAGCCA GTCTGCAGGT CGACCATAGT GACTGGATAT
3001 GTTGTGTTTT ACAGTATTAT GTAGTCTGTT TTTTATGCAA AATCTAATTT AATATATTGA
3061 TATTTATATC ATTTTACGTT TCTCGTTCAG CTTTCTTGTA CAAAGTGGTT TGATGGCCGC
3121 TAAGTAAGTA AGACGTCGAG CTCTAAGTAA GTAACGGCCG CCACCGCGGT GGAGCTTTGG
3181 ACTTCTTCGC CAGAGGTTTG GTCAAGTCTC CAATCAAGGT TGTCGGCTTG TCTACCTTGC
3241 CAGAAATTTA CGAAAAGATG GAAAAGGGTC AAATCGTTGG TAGATACGTT GTTGACACTT
3301 CTAAATAAGC GAATTTCTTA TGATTTATGA TTTTATTAT TAAATAAGTT ATAAAAAAA
3361 TAAGTGTATA CAAATTTTAA AGTGACTCTT AGGTTTTTAA ACGAAAATTC TTATTCTTGA
3421 GTAACCTTTT CCTGTAGGTC AGGTGCTTT CTCAGGTATA GCATGAGGTC GCTCTTATTG
3481 ACCACACCTC TACCGCATG CCGAGCAAAT GCCTGCAAAT CGCTCCCAT TTCACCCAAT
3541 TGTAGATATG CTAAGTCCAG CAATGAGTTG ATGAATCTCG GTGTGTATTT TATGTCTCA
3601 GAGGACAATA CCTGTTGTAA TCGTTCTTCC ACACGGATCC CAATTCGCCC TATAGTGAGT
3661 CGTATTACAA TTCACTGGCC GTCGTTTTAC AACGTCGTGA CTGGGAAAAC CCTGGCGTTA
3721 CCCAACTTAA TCGCCTTGCA GCACATCCCC CTTTCGCCAG CTGGCGTAAT AGCGAAGAGG
3781 CCCGCACCGA TCGCCCTTCC CAACAGTTGC GCAGCCTGAA TGGCGAATGG ACGCGCCCTG
3841 TAGCGGCGCA TTAAGCGCGG CGGGTGTGGT GGTACGCGC AGCGTGACCG CTACACTTGC
3901 CAGCGCCCTA GCGCCCGCTC CTTTCGCTTT CTTCCCTTCC TTTCTCGCCA CGTTCGCGG
3961 CTTTCCCCGT CAAGCTCTAA ATCGGGGGCT CCCTTTAGGG TTCCGATTTA GTGCTTTACG
4021 GCACCTCGAC CCCAAAAAAC TTGATTAGGG TGATGGTTCA CGTAGTGGGC CATCGCCCTG
4081 ATAGACGGTT TTTCCGCCCT TGACGTTGGA GTCCACGTTT TTTAATAGTG GACTCTTGTT
4141 CCAAAGTGA ACAACACTCA ACCCTATCTC GGTCTATTCT TTTGATTAT AAGGGATTTT
4201 GCCGATTTTC GCCTATTGGT TAAAAAATGA GCTGATTTAA CAAAAATTTA ACGCGAATTT
4261 TAACAAAATA TTAACGTTTA CAATTTCTTG ATGCGGTATT TTCTCCTTAC GCATCTGTGC
4321 GGTATTTTCA ACCGCAGGCA AGTGACAAA CAATACTTAA ATAAATACTA CTCAGTAATA
4381 ACCTATTTCT TAGCATTTT GACGAAATTT GCTATTTTGT TAGAGTCTTT TACACATTT
4441 GTCTCCACAC CTCCGCTTAC ATCAACACCA ATAACGCCAT TTAATCTAAG CGCATCACCA
4501 ACATTTTCTG GCGTCAGTCC ACCAGCTAAC ATAAATGTA AGCTTTCCGG GCTCTCTTGC
4561 CTTCCAACCC AGTCAGAAAT CGAGTTCCAA TCCAAAAGTT CACCTGTCCC ACCTGCTTCT
4621 GAATCAAACA AGGGAATAAA CGAATGAGGT TTCTGTGAAG CTGCACTGAG TAGTATGTTG
4681 CAGTCTTTTG GAAATACGAG TCTTTTAATA ACTGGCAAAC CGAGGAATCT TTGGTATTCT
4741 TGCCACGACT CATCTCCATG CAGTTGGACG ATATCAATGC CGTAATCATT GACCAGAGCC
4801 AAAACATCCT CCTTAGGTTG ATTACGAAAC ACGCCAACCA AGTATTTCCG AGTGCCGTGA
4861 CTATTTTAT ATGCTTTTAC AAGACTTGAA ATTTTCTTGG CAATAACCGG GTCAATTGTT
4921 CTCTTTCTAT TGGGCACACA TATAATACCC AGCAAGTCAG CATCGGAATC TAGAGCACAT
4981 TCTGCGGCT CTGTGCTCTG CAAGCCGCAA ACTTTCACCA ATGGACCAGA ACTACCTGTG
5041 AAATTAATAA CAGACATACT CCAAGCTGCC TTTGTGTGCT TAATCACGTA TACTCACGTG
5101 CTCAATAGTC ACCAATGCCC TCCCTCTTGG CCCTCTCCTT TTCTTTTTTC GACCGAATTA
5161 ATTCTTAATC GGCAAAAAA GAAAAGCTCC GGATCAAGAT TGTACGTAAG GTGACAAGCT
5221 ATTTTTCAT AAAGAATATC TTCCACTACT GCCATCTGGC GTCATAACTG CAAAGTACAC
5281 ATATATTACG ATGCTGTCTA TTAATGCTT CCTATATTAT ATATATAGTA ATGTCTGTTA
5341 TGGTGCACTC TCAGTACAAT CTGCTCTGAT GCCGCATAGT TAAGCCAGCC CCGACACCCG
5401 CCAACACCCG CTGACGCGCC CTGACGGGCT TGTCTGCTCC CGGCATCCGC TTACAGACAA
5461 GCTGTGACCG TCTCCGGGAG CTGCATGTGT CAGAGGTTTT CACCGTCATC ACCGAAACGC
5521 GCGAGACGAA AGGGCCTCGT GATACGCTA TTTTATAGG TTAATGTCAT GATAATAATG
5581 GTTCTTAGG ACGGATCGCT TGCCTGTAAC TTACACGCGC CTCGTATCTT TTAATGATGG
5641 AATAATTTGG GAATTTACTC TGTGTTTATG TATTTTATG TTTTGTATTT GGATTTTGA
5701 AAGTAAATAA AGAAGGTAGA AGAGTTACGG AATGAAGAAA AAAAAATAAA CAAAGTTTAA
5761 AAAAATTTCA ACAAAAAGCG TACTTTACAT ATATATTTAT TAGACAAGAA AAGCAGATTA
5821 AATAGATATA CATTTCGATTA ACGATAAGTA AAATGTAAAA TCACAGGATT TTCGTGTGTG
5881 GTCTTCTACA CAGACAAGAT GAAACAATTC GGCATTAATA CCTGAGAGCA GGAAGAGCAA
5941 GATAAAAGGT AGTATTTGTT GCGCATCCCC CTAGAGTCTT TTACATCTTC GGAAAACAAA
6001 AACTATTTTT TCTTTAATTT CTTTTTTTAC TTTCTATTTT TAATTTATAT ATTTATATTA-

FIGURE 42C

119/240

6061 AAAAATTTAA ATTATAATTA TTTTATAGC ACGTGATGAA AAGGACCCAG GTGGCACTTT
6121 TCGGGGAAAT GTGCGCGGAA CCCCTATTTG TTTATTTTTC TAAATACATT CAAATATGTA
6181 TCCGCTCATG AGACAATAAC CCTGATAAAT GCTTCAATAA TATTGAAAAA GGAAGAGTAT
6241 GAGTATTCAA CATTTCCGTG TCGCCCTTAT TCCCTTTTTT GCGGCATTTT GCCTTCCTGT
6301 TTTTGCTCAC CCAGAAACGC TGGTGAAAGT AAAAGATGCT GAAGATCAGT TGGGTGCACG
6361 AGTGGGTTAC ATCGAACTGG ATCTCAACAG CGGTAAGATC CTTGAGAGTT TTCGCCCCGA
6421 AGAACGTTTT CCAATGATGA GCACTTTTAA AGTTCTGCTA TGTGGCGCGG TATTATCCCG
6481 TATTGACGCC GGGCAAGAGC AACTCGGTCT CCGCATACAC TATTCTCAGA ATGACTTGGT
6541 TGAGTACTCA CCAGTCACAG AAAAGCATCT TACGGATGGC ATGACAGTAA GAGAATTATG
6601 CAGTGCTGCC ATAACCATGA GTGATAACAC TGCGGCCAAC TTACTTCTGA CAACGATCGG
6661 AGGACCGAAG GAGCTAACCG CTTTTTTTCA CAACATGGGG GATCATGTAA CTCGCCTTGA
6721 TCGTTGGGAA CCGGAGCTGA ATGAAGCCAT ACCAAACGAC GAGCGTGACA CCACGATGCC
6781 TGTAGCAATG GCAACAACGT TGCGCAACT ATTAACCTGGC GAACTACTTA CTCTAGCTTC
6841 CCGGCAACAA TTAATAGACT GGATGGAGGC GGATAAAGTT GCAGGACCAC TTCTGCGCTC
6901 GGCCCTTCCG GCTGGCTGGT TTATTGCTGA TAAATCTGGA GCCGGTGAGC GTGGGTCTCG
6961 CGGTATCATT GCAGCACTGG GGCCAGATGG TAAGCCCTCC CGTATCGTAG TTATCTACAC
7021 GACGGGCAGT CAGGCAACTA TGGATGAACG AAATAGACAG ATCGCTGAGA TAGGTGCCTC
7081 ACTGATTAAAG CATTGGTAAC TGTCAGACCA AGTTTACTCA TATATACTTT AGATTGATTT
7141 AAAACTTCAT TTTTAATTTA AAAGGATCTA GGTGAAGATC CTTTTTGATA ATCTCATGAC
7201 CAAAATCCCT TAACGTGAGT TTTCTGTCCA CTGAGCGTCA GACCCCGTAG AAAAGATCAA
7261 AGGATCTTCT TGAGATCCTT TTTTCTGCG CGTAATCTGC TGCTTGCAAA CAAAAAACC
7321 ACCGCTACCA GCGGTGGTTT GTTTGCCGGA TCAAGAGCTA CCAACTCTTT TTCCGAAGGT
7381 AACTGGCTTC AGCAGAGCGC AGATACCAAA TACTGCTCTT CTAGTGTAGC CGTAGTTAGG
7441 CCACCACTTC AAGAACTCTG TAGCACCGCC TACATACCTC GCTCTGCTAA TCCTGTTACC
7501 AGTGGCTGCT GCCAGTGGCG ATAAGTCGTG TCTTACCGGG TTGGACTCAA GACGATAGTT
7561 ACCGGATAAG GCGCAGCGGT CGGGCTGAAC GGGGGGTTCC TGACACACAGC CCAGCTTGGA
7621 GCGAACGACC TACACCGAAC TGAGATACCT ACAGCGTGAG CATTGAGAAA GCGCCACGCT
7681 TCCCGAAGGG AGAAAGGCGG ACAGGTATCC GGTAAAGCGC AGGGTCGGAA CAGGAGAGCG
7741 CACGAGGGAG CTTCCAGGGG GGAACGCCTG GTATCTTTAT AGTCCTGTGC GGTTCGCCA
7801 CCTCTGACTT GAGCGTCGAT TTTTGTGATG CTCGTACAGG GGGCCGAGCC TATGGA AAAA
7861 CGCCAGCAAC GCGGCCTTTT TACGGTTCTT GGCCTTTTGC TGGCCTTTTG CTCACATGTT
7921 CTTTCTGTGG TTATCCCCTG ATTCTGTGGA TAACCGTATT ACCGCCTTTG AGTGAGCTGA
7981 TACCGCTCGC CGCAGCCGAA CGACCGAGCG CAGCGAGTCA GTGAGCGAGG AAGCGGAAGA
8041 GCGCCCAATA CGCAAACCGC CTCTCCCCGC GCGTTGGCCG ATTCATTAAT GCAGCTGGCA
8101 CGACAGGTTT CCCGACTGGA AAGCGGGCAG TGAGCGCAAC GCAATTAATG TGAGTTACCT
8161 CACTCATTAG GCACCCCAAG CTTTACACTT TATGCTTCCG GCTCCTATGT TGTGTGGAAT
8221 TGTGAGCGGA TAACAATTTT ACACAGGAAA CAGCTATGAC CATGATTACG CCAAGCTCGG
8281 AATTAACCTT CACTAAAGGG AACAAAAGCT GGGTACCGGG CCCCCCTCG AGATCCGGGA
8341 TCGAAGAAAT GATGGTAAAT GAAATAGGAA ATCAAGGAGC ATGAAGGCAA AAGACAAATA
8401 TAAGGGTCGA ACGAAAAATA AAGTGAAAG TGTGATATG ATGTATTTGG CTTTGC GGCG
8461 CCGAAAAAAC GAGTTTACGC AATTGCACAA TCATGCTGAC TCTGTGGCGG ACCCGCGCTC
8521 TTGCCGGCCC GCGGATAACG CTGGGCGTGA GGCTGTGCCC GGCGGAGTTT TTTGCGCCTG
8581 CATTTTCCAA GGTTTACCCT GCGCTAAGGG GCGAGATTGG AGAAGCAATA AGAATGCCGG
8641 TTGGGGTTGC GATGATGACG ACCACGACAA CTGGTGTCTAT TATTTAAGTT GCCGAAAGAA
8701 CCTGAGTGCA TTTGCAACAT GAGTATACTA GAAGAATGAG CCAAGACTTG CGAGACGCGA
8761 GTTTGCCGGT GGTGCGAACA ATAGAGCGAC CATGACCTTG AAGGTGAGAC GCGCATAACC
8821 GCTAGAGTAC TTTGAAGAGG AAACAGCAAT AGGGTTGCTA CCAGTATAAA TAGACAGGTA
8881 CATAACAAC TGGAAATGGT TGTCTGTTTG AGTACGCTTT CAA

FIGURE 42D

120/240

pDEST23

His6 carboxy-fusion vector, T7 promoter

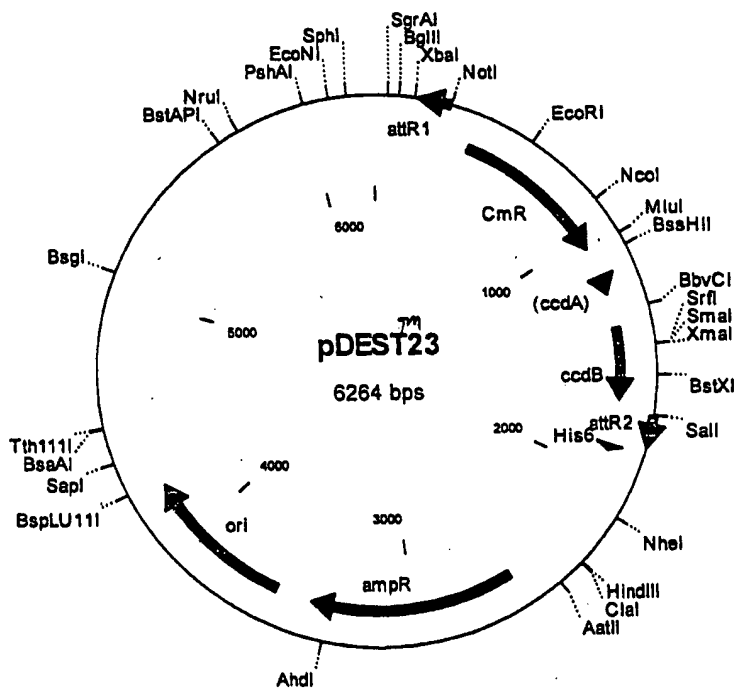
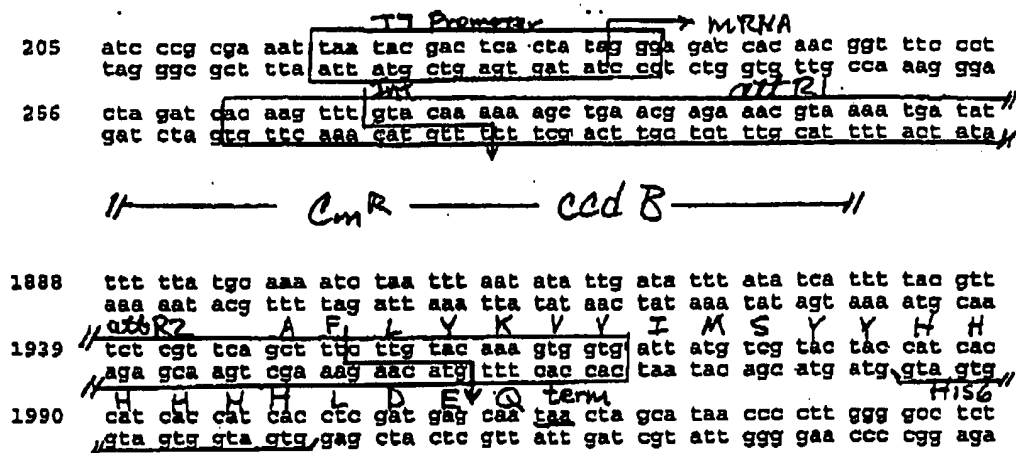


FIGURE 43A

pDEST23 6264 bp

<u>Location (Base Nos.)</u>		<u>Gene Encoded</u>
285..161		attR1
394..1053		CmR
1173..1257		inactivated ccdA
1395..1700		ccdB
1741..1865		attR2
1883..1911		his6
2574..3434		ampR
3583..4222		ori

1	TCCTCCCCAT	CGGTGATGTC	GGCGATATAG	GCGCCAGCAA	CCGCACCTGT	GGCGCCGGTG
61	ATGCCGGCCA	CGATGCGTCC	GGCGTAGAGG	ATCGAGATCT	CGATCCCCGG	AAATTAATAC
121	GACTCACTAT	AGGGAGACCA	CAACGGTTTC	CCTCTAGATC	ACAAGTTTGT	ACAAAAAAGC
181	TGAACGAGAA	ACGTAAATG	ATATAAATAT	CAATATATTA	AATTAGATT	TGCATAAAAA
241	ACAGACTACA	TAATACTGTA	AAACACAACA	TATCCAGTCA	CTATGGCGGC	CGCATTAGGC
301	ACCCCGAGGCT	TTACACTTTA	TGCTTCCGGC	TCGTATAATG	TGTGGATT	GAGTTAGGAT
361	CCGGCGAGAT	TTTCAGGAGC	TAAGGAAGCT	AAAATGGAGA	AAAAAATCAC	TGGATATACC
421	ACCGTTGATA	TATCCCAATG	GCATCGTAAA	GAACATTTTG	AGGCATTTC	GTGAGTTGCT
481	CAATGTACCT	ATAACCGAGC	CGTTCAGCTG	GATATTACGG	CCTTTTTTAA	GACCGTAAAG
541	AAAAATAAGC	ACAAGTTT	TCCGGCCTTT	ATTCACATTC	TTGCCCGCCT	GATGAATGCT
601	CATCCGGAAT	TCCGTATGGC	AATGAAAGAC	GGTGAGCTGG	TGATATGGGA	TAGTGTTCAC
661	CCTTGTTACA	CCGTTTTCCA	TGAGCAAAC	GAAACGTTTT	CATCGCTCTG	GAGTGAATAC
721	CACGACGATT	TCCGGCAGTT	TCTACACATA	TATTGCAAG	ATGTGGCGTG	TTACGGTGAA
781	AACCTGGCCT	ATTTCCCTAA	AGGGTTTATT	GAGAATATGT	TTTTCGTCTC	AGCCAATCCC
841	TGGGTGAGTT	TCACCAGTTT	TGATTTAAAC	GTGGCCAATA	TGGACAACCT	CTTCGCCCCC
901	GTTTTACCA	TGGGCAAATA	TTATACGCAA	GGCGACAAGG	TGCTGATGCT	GCTGGCGATT
961	CAGGTTTCATC	ATGCCGTCTG	TGATGGCTTC	CATGTCGGCA	GAATGCTTAA	TGAATTACAA
1021	CAGTACTGCG	ATGAGTGCGA	GGGCGGGGCG	TAAACGCGTG	GATCCGGCTT	ACTAAAAGCC
1081	AGATAACAGT	ATGCGTATTT	GCGCGCTGAT	TTTTGCGGTA	TAAGAATATA	TACTGATATG
1141	TATACCCGAA	GTATGTCAAA	AAGAGGTGTG	CTATGAAGCA	GCGTATTACA	GTGACAGTTG
1201	ACAGCGACAG	CTATCAGTTG	CTCAAGGCAT	ATATGATGTC	AATATCTCCG	GTCTGGTAAG
1261	CACAACCATG	CAGAATGAAG	CCCGTCGTCT	GCGTGCCGAA	CGCTGGAAG	CGGAAAATCA
1321	GGAAGGGATG	GCTGAGGTCG	CCCGGTTTAT	TGAAATGAAC	GGCTCTTTTG	CTGACGAGAA
1381	CAGGGAACTGG	TGAAATGCAG	TTTAAGGTTT	ACACCTATAA	AAGAGAGAGC	CGTTATCGTC
1441	TGTTTGTGGA	TGTACAGAGT	GATATTATTG	ACACGCCCGG	GCGACGGATG	GTGATCCCCC
1501	TGGCCAGTGC	ACGTCTGCTG	TCAGATAAAG	TCTCCCGTGA	ACTTTACCCG	GTGGTGCATA
1561	TCGGGGATGA	AAGCTGGCGC	ATGATGACCA	CCGATATGGC	CAGTGTGCCG	GTCTCCGTTA
1621	TCGGGGAAGA	AGTGGCTGAT	CTCAGCCACC	GCGAAAATGA	CATCAAAAAC	GCCATTAACC
1681	TGATGTTCTG	GGGAATATAA	ATGTCAGGCT	CCCTTATACA	CAGCCAGTCT	GCAGGTCGAC
1741	CATAGTGA	GGATATGTTG	TGTTTTACAG	TATTATGTAG	TCTGTTTTTT	ATGCAAAATC
1801	TAATTTAATA	TATTGATATT	TATATCATTT	TACGTTTCTC	GTTGATGATA	CTGTTGATCA
1861	GTGGTGATTA	TGTCGTACTA	CCATCACC	CACCATCACC	TCGATGAGCA	ATAACTAGCA
1921	TAACCCCTTG	GGGCCTCTAA	ACGGGTCTTG	AGGGGTTTTT	TGCTGAAAGG	AGGAACTATA
1981	TCCGGATATC	CACAGGACGG	GTGTGGTCGC	CATGATCGCG	TAGTCGATAG	TGGCTCCAAG
2041	TAGCGAAGCG	AGCAGGACTG	GGCGGCGGCC	AAAGCGGTCG	GACAGTGCTC	CGAGAACGGG
2101	TGCGCATAGA	AATTGCATCA	ACGCATATAG	CGCTAGCAGC	ACGCCATAGT	GACTGGCGAT
2161	GCTGTCGGAA	TGGACGATAT	CCCGCAAGAG	GCCCCGCGAGT	ACCGGCATAA	CCAAGCCTAT
2221	GCCTACAGCA	TCCAGGGTGA	CGGTGCCGAG	GATGACGATG	AGCGCATTGT	TAGATTTTAT
2281	ACACGGTGCC	TGACTGCGTT	AGCAATTFAA	CTGTGATAAA	CTACCGCATT	AAAGCTTATC
2341	GATGATAAGC	TGTCAAACAT	GAGAATTCCT	GAAGACGAAA	GGGCCTCGTG	ATACGCCTAT
2401	TTTTATAGGT	TAATGTCATG	ATAATAATGG	TTTCTTAGAC	GTCAGGTGGC	ACTTTTCGGG
2461	GAAATGTGCG	CGGAACCCCT	ATTTGTTTAT	TTTTCTAAAT	ACATTCAAAT	ATGTATCCGC
2521	TCATGAGACA	ATAACCCCTGA	TAAATGCTTC	AATAATATTG	AAAAAGGAAG	AGTATGAGTA
2581	TTCAACATTT	CCGTGTCGCC	CTTATTCCTT	TTTTTGCGGC	ATTTTGCTCT	CCTGTTTTTG
2641	CTCACCCAGA	AACGCTGGTG	AAAGTAAAAG	ATGCTGAAGA	TCAGTTGGGT	GCACGAGTGG

FIGURE 43B

122/240

2701 GTTACATCGA ACTGGATCTC AACAGCGGTA AGATCCTTGA GAGTTTTCGC CCCGAAGAAC
2761 GTTTTCCAAT GATGAGCACT TTTAAAGTTC TGCTATGTGG CGCGGTATTA TCCCGTGTG
2821 ACGCCGGGCA AGAGCAACTC GGTCCGCCGA TACACTATTC TCAGAATGAC TTGGTTGAGT
2881 ACTCACCAGT CACAGAAAAG CATCTTACGG ATGGCATGAC AGTAAGAGAA TTATGCAGTG
2941 CTGCCATAAC CATGAGTGAT AACACTGCGG CCAACTTACT TCTGACAAAG ATCGGAGGAC
3001 CGAAGGAGCT AACCGCTTTT TTGCACAACA TGGGGGATCA TGTAACCTCGC CTTGATCGTT
3061 GGGAAACCGGA GCTGAATGAA GCCATACCAA ACGACGAGCG TGACACCACG ATGCCTGCAG
3121 CAATGGCAAC AACGTTGCGC AAACCTATTAA CTGGCGAACT ACTTACTCTA GCTTCCCGGC
3181 AACAAATTAAT AGACTGGATG GAGGCGGATA AAGTTGCAGG ACCACTTCTG CGCTCGGCCC
3241 TTCCGGCTGG CTGGTTTATT GCTGATAAAT CTGGAGCCGG TGAGCGTGGG TCTCGCGGTA
3301 TCATTGCAGC ACTGGGGCCA GATGGTAAGC CCTCCCGTAT CGTAGTTATC TACACGACGG
3361 GGAGTCAGGC AACTATGGAT GAACGAAATA GACAGATCGC TGAGATAGGT GCCTCACTGA
3421 TTAAGCATTG GTAAGTGTC GACCAAGTTT ACTCATATAT ACTTTAGATT GATTTAAAGC
3481 TTCATTTTTA ATTTAAAGG ATCTAGGTGA AGATCCTTTT TGATAATCTC ATGACCAAAA
3541 TCCCTTAACG TGAGTTTTCG TTCCACTGAG CGTCAGACCC CGTAGAAAAG ATCAAAGGAT
3601 CTTCTTGAGA TCCTTTTTTT CTGCGCGTAA TCTGTGCTT GCAAACAAA AAACCACCGC
3661 TACCAGCGGT GGTGTGTTT CCGGATCAAG AGCTACCAAC TCTTTTTCCG AAGGTAACCTG
3721 GCTTCAGCAG AGCGCAGATA CCAAATACTG TCCTTCTAGT GTAGCCGTAG TTAGGCCACC
3781 ACTTCAAGAA CTCTGTAGCA CCGCTACAT ACCTCGCTCT GCTAATCCTG TTACCACTGG
3841 CTGCTGCCAG TGGCGATAAG TCGTGTCTTA CCGGTTGGA CTCAAGACGA TAGTTACCGG
3901 ATAAGGCGCA GCGGTCGGGC TGAACGGGGG GTTCGTGCAC ACAGCCCAGC TTGGAGCGAA
3961 CGACCTACAC CGAACTGAGA TACCTACAGC GTGAGCTATG AGAAAGCGCC ACGCTTCCCG
4021 AAGGGAGAAA GGCGGACAGG TATCCGTTAA GCGGCAGGGT CGGAACAGGA GAGCGCACGA
4081 GGGAGCTTCC AGGGGGAAAC GCCTGGTATC TTTATAGTCC TGTCGGGTTT CGCCACCTCT
4141 GACTTGAGCG TCGATTTTTG TGATGCTCGT CAGGGGGGCG GAGCCTATGG AAAAACGCCA
4201 GCAACGCGGC CTTTTTACGG TTCCTGGCCT TTTGCTGGCC TTTTGCTCAC ATGTTCTTTC
4261 CTGCGTTATC CCCTGATTCT GTGGATAACC GTATTACCGC CTTTGAGTGA GCTGATACCG
4321 CTCGCCGAG CCGAACGACC GAGCGCAGCG AGTCAGTGAG CGAGGAAGCG GAAGAGCGCC
4381 TGATGCGGTA TTTTCTCCTT ACGCATCTGT GCGGTATTTT ACACCGCATA TATGGTGCAC
4441 TCTCAGTACA ATCTGCTCTG ATGCCGCATA GTTAAGCCAG TATACACTCC GCTATCGCTA
4501 CGTGACTGGG TCATGGCTGC GCCCCGACAC CCGCCAACAC CCGCTGACGC GCCCTGACGG
4561 GCTTGTCTGC TCCCGGCATC CGCTTACAGA CAAGCTGTGA CCGTCTCCGG GAGCTGCATG
4621 TGTCAGAGGT TTTACCGTC ATCACCAGAA CGCGCGAGGC AGCTGCGGTA AAGCTCATCA
4681 GCGTGGTCTG GAAGCGATT CACAGATGTCT GCCTGTTTAT CCGCGTCCAG CTCGTTGAGT
4741 TTCTCCAGAA GCGTTAATGT CTGGCTTCTG ATAAAGCGGG CCATGTTAAG GCGGTTTCTT
4801 TCCTGTTTGG TCACTGATGC CTCCGTGTAA GGGGGATTTT TGTTTATGGG GGTAATGATA
4861 CCGATGAAAC GAGAGAGGAT GCTCACGATA CCGGTTACTG ATGATGAACA TGCCCGGTTA
4921 CTGGAACGTT GTGAGGGTAA ACAACTGGCG GTATGGATGC GCGGGGACCA GAGAAAAATC
4981 ACTCAGGGTC AATGCCAGCG CTTTCGTTAAT ACAGATGTAG GTGTTCCACA GGGTAGCCAG
5041 CAGCATCCTG CGATGCAGAT CCGGAACATA ATGGTGCAGG GCGCTGACTT CCGCGTTTCC
5101 AGACTTTACG AAACACGGAA ACCGAAGACC ATTCATGTTG TTGCTCAGGT CGCAGACGTT
5161 TTGCAGCAGC AGTCGCTTCA CGTTCGCTCG CGTATCGGTG ATTCATTCTG CTAACCAGTA
5221 AGGCAACCCC GCCAGCCTAG CCGGGTCTCT AACGACAGGA GCACGATCAT GCGCACCCGT
5281 GGCCAGGACC CAACGCTGCC CGAGATGCGC CGCGTGCGGC TGCTGGAGAT GCGGAGCGCG
5341 ATGGATATGT TCTGCCAAGG GTTGGTTTGC GCATTACAG TTCTCCGCAA GAATTGATTG
5401 GCTCCAATTC TTGGAGTGGT GAATCCGTTA GCGAGGTGCC GCCGGCTTCC ATTCAGGTG
5461 AGGTGGCCCG GCTCCATGCA CCGCGACGCA ACGCGGGGAG GCAGACAAGG TATAGGGCGG
5521 CGCTACAAT CCATGCCAAC CCGTTCCATG TGCTCGCCGA GCGGGCATAA ATCGCCGTGA
5581 CGATCAGCGG TCCAGTGATC GAAGTTAGGC TGTAAGAGC CGCGAGCGAT CCTTGAAGCT
5641 GTCCCTGATG GTCGTCATCT ACCTGCCTGG ACAGCATGGC CTGCAACGCG GGTATCCGGA
5701 TGCCGCCGGA AGCGAGAAGA ATCATAATGG GGAAGGCCAT CCAGCCTCGC GTCGCGAAGC
5761 CCAGCAAGAC GTAGCCAGC GCGTCGGCCG CCATGCCGCG GATAATGGCC TGCTTCTCGC
5821 CGAAACGTTT GGTGGCGGGA CCAGTGACGA AGGCTTGAGC GAGGGCGTGC AAGATTCCGA
5881 ATACCGCAAG CGACAGGCCG ATCATCGTCC CGCTCCAGCG AAAGCGGTCC TCGCCGAAAA
5941 TGACCCAGAG CGCTGCCGGC ACCTGTCTTA CGAGTTGCAT GATAAAGAAG ACAGTCATAA
6001 GTGCGGCGAC GATAGTCATG CCCCAGCGCC ACCGGAAGGA GCTGACTGGG TTGAAGGCTC
6061 TCAAGGGCAT CGGTGCATCG ACGCTCTCCC TTATGCGACT CCTGCATTAG GAAGCAGCCC
6121 AGTAGTAGGT TGAGGCCGTT GAGCACCACC GCCGCAAGGA ATGGTGCATG CAAGGAGATG -

FIGURE 43C

123/240

6181 GCGCCCAACA GTCCCCCGGC CACGGGGCCT GCCACCATAC CCACGCCGAA ACAAGCGCTC
6241 ATGAGCCCGA AGTGGCGAGC CCGA

FIGURE 43D

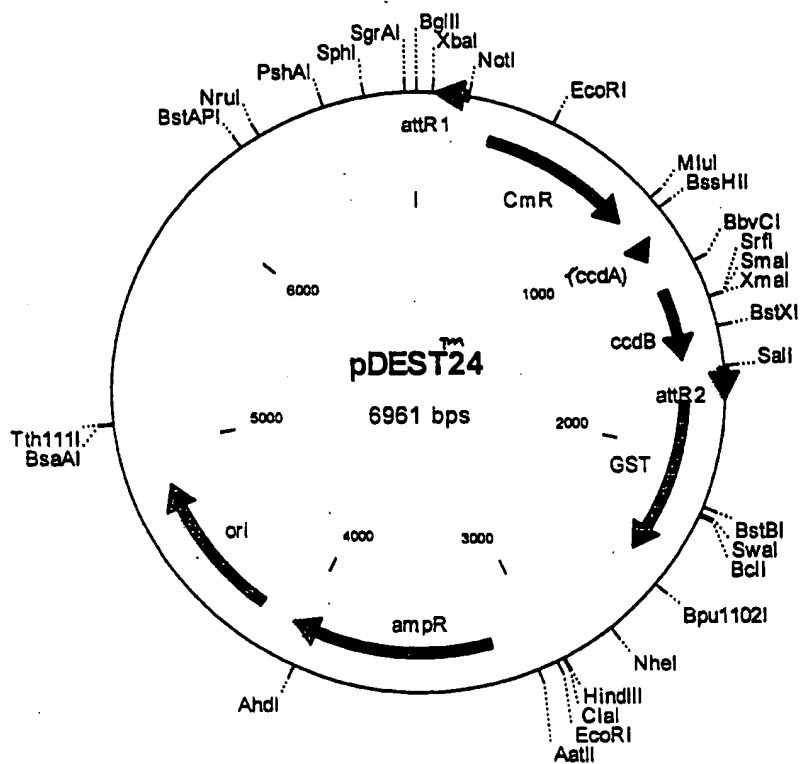
[illegible]

FIGURE 44A

125/240

pDEST24 6961 bp

<u>Location (Base Nos.)</u>		<u>Gene Encoded</u>
195..71		attR1
304..963		CmR
1083..1167		inactivated ccdA
1305..1610		ccdB
1651..1775		attR2
1783..2451		GST
3181..4041		ampR
4190..4829		ori

1	ATCGAGATCT	CGATCCCGCG	AAATTAATAC	GACTCACTAT	AGGGAGACCA	CAACGGTTTC
61	CCTCTAGATC	ACAAGTTTGT	ACAAAAAGC	TGAACGAGAA	ACGTAAAATG	ATATAAATAT
121	CAATATATTA	AATTAGATTT	TGCATAAAAA	ACAGACTACA	TAATACTGTA	AAACACAACA
181	TATCCAGTCA	CTATGGCGGC	CGCATTAGGC	ACCCCAGGCT	TTACTACTTA	TGCTTCCGGC
241	TCGTATAATG	TGTGGATTTT	GAGTTAGGAT	CCGGCGAGAT	TTTCAGGAGC	TAAGGAAGCT
301	AAAATGGAGA	AAAAAATCAC	TGGATATACC	ACCGTTGATA	TATCCCAATG	GCATCGTAAA
361	GAACATTTTG	AGGCATTTCA	GTCACTTGCT	CAATGTACCT	ATAACCAGAC	CGTTCAGCTG
421	GATATTACGG	CCTTTTTTAA	GACCGTAAAG	AAAAATAAGC	ACAAGTTTTA	TCCGGCCTTT
481	ATTCACATTC	TTGCCCGCCT	GATGAATGCT	CATCCGGAAT	TCCGTATGGC	AATGAAAGAC
541	GGTGAGCTGG	TGATATGGGA	TAGTGTTCAC	CCTTGTTACA	CCGTTTTCCA	TGAGCAAAC
601	GAAACGTTTT	CATCGCTCTG	GAGTGAATAC	CACGACGATT	TCCGGCAGTT	TCTACACATA
661	TATTCGCAAG	ATGTGGCGTG	TTACGGTGAA	AACCTGGCCT	ATTTCCCTAA	AGGGTTTATT
721	GAGAATATGT	TTTTCGTCTC	AGCCAAATCCC	TGGGTGAGTT	TCACCAGTTT	TGATTTAAAC
781	GTGGCCAATA	TGGACAACCT	CTTCGCCCCC	GTTTTACCA	TGGGCAAATA	TTATACGCAA
841	GGCGACAAGG	TGCTGATGCC	GCTGGCGATT	CAGGTTTCATC	ATGCCGTCTG	TGATGGCTTC
901	CATGTCGGCA	GAATGCTTAA	TGAATTACAA	CAGTACTGCG	ATGAGTGGCA	GGCGGGGGCG
961	TAAACGCGTG	GATCCGGCTT	ACTAAAAGCC	AGATAACAGT	ATGCGTATTT	GCGCGCTGAT
1021	TTTTGCGGTA	TAAGAATATA	TACTGATATG	TATACCCGAA	GTATGTCAAA	AAGAGGTGTG
1081	CTATGAAGCA	GCGTATTACA	GTGACAGTTG	ACAGCGACAG	CTATCAGTTG	CTCAAGGCAT
1141	ATATGATGTC	AATATCTCCG	GTCTGGTAAG	CACAACCATG	CAGAATGAAG	CCCGTCGTCT
1201	GCGTGCCGAA	CGCTGGAAAG	CGGAAAATCA	GGAAGGGATG	GCTGAGGTCTG	CCCGGTTTAT
1261	TGAAATGAAC	GGCTCTTTTG	CTGACGAGAA	CAGGGACTGG	TGAAATGCAG	TTTAAGGTTT
1321	ACACCTATAA	AAGAGAGAGC	CGTTATCGTC	TGTTTGTGGA	TGTACAGAGT	GATATTATTG
1381	ACACGCCCCG	GCGACGGATG	GTGATCCCCC	TGGCCAGTGC	ACGTCTGCTG	TCAGATAAAG
1441	TCTCCCGTGA	ACTTTACCCG	GTGGTGCATA	TCGGGGATGA	AAGCTGGCGC	ATGATGACCA
1501	CCGATATGGC	CAGTGTGCCG	GTCTCCGTTA	TCGGGGAAGA	AGTGGCTGAT	CTCAGCCACC
1561	GCGAAAATGA	CATCAAAAAC	GCCATTAAAC	TGATGTTCTG	GGGAATATAA	ATGTCAGGCT
1621	CCCTTATACA	CAGCCAGTCT	GCAGGTCGAC	CATAGTGAAT	GGATATGTTG	TGTTTTACAG
1681	TATTATGTAG	TCTGTTTTTT	ATGCAAAATC	TAATTTAATA	TATTGATATT	TATATCATT
1741	TACGTTTCTC	GTTTCAGCTT	CTTGATACAA	GTGGTGATTA	TGTCCCTAT	ACTAGGTTAT
1801	TGGAAAATTA	AGGGCCTTGT	GCAACCCACT	CGACTTCTTT	TGGAATATCT	TGAAGAAAAA
1861	TATGAAGAGC	ATTGTATGA	GCGCGATGAA	GGTGATAAAT	GGCGAAACAA	AAAGTTTGAA
1921	TTGGGTTTGG	AGTTTCCCAA	TCTTCCTTAT	TATATTGATG	GTGATGTTAA	ATTAACACAG
1981	TCTATGGCCA	TCATACGTTA	TATAGCTGAC	AAGCACAACA	TGTTGGGTGG	TTGTCCAAAA
2041	GAGCGTGCG	AGATTTCAAT	GCTTGAAGGA	GCGGTTTTGG	ATATTAGATA	CGGTGTTTCG
2101	AGAATTGCAT	ATAGTAAAGA	CTTTGAAACT	CTCAAAGTTG	ATTTTCTTAG	CAAGCTACCT
2161	GAAATGCTGA	AAATGTTCTG	AGATCGTTTA	TGTCATAAAA	CATATTTAAA	TGGTGATCAT
2221	GTAACCCATC	CTGACTTCAT	GTTGTATGAC	GCTCTTGATG	TTGTTTTATA	CATGGACCCA
2281	ATGTGCCTGG	ATGCGTTCCC	AAAATTAGTT	TGTTTTAAAA	AACGTATTGA	AGCTATCCCA
2341	CAAATTGATA	AGTACTTGAA	ATCCAGCAAG	TATATAGCAT	GGCCTTTGCA	GGGCTGGCAA
2401	GCCACGTTTG	GTGGTGGCGA	CCATCCTCCA	AAATCGGATC	TGGTTCCGCG	TCCATGGGGA
2461	TCCGGCTGCT	AACAAAGCCC	GAAAGGAAGC	TGAGTTGGCT	GCTGCCACCG	CTGAGCAATA
2521	ACTAGCATAA	CCCCTTGGGG	CCTCTAAACG	GGTCTTGAGG	GGTTTTTTGC	TGAAAGGAGG
2581	AACTATATCC	GGATATCCAC	AGGACGGGTG	TGGTCGCCAT	GATCGCGTAG	TCGATAGTGG
2641	CTCCAAGTAG	CGAAGCGAGC	AGGACTGGGC	GGCGGCCAAA	GCGGTCGGAC	AGTGCTCCGA-

FIGURE 44B

126/240

```

2701 GAACGGGTGC GCATAGAAAT TGCATCAACG CATATAGCGC TAGCAGCAGC CCATAGTGAC
2761 TGGCGATGCT GTCGGAATGG ACGATATCCC GCAAGAGGCC CGGCAGTACC GGCATAACCA
2821 AGCCTATGCC TACAGCATCC AGGGTGACGG TGCCGAGGAT GACGATGAGC GCATTGTTAG
2881 ATTTCATACA CGGTGCCTGA CTGCGTTAGC AATTTAACTG TGATAAACTA CCGCATTAAA
2941 GCTTATCGAT GATAAGCTGT CAAACATGAG AATTCTTGAA GACGAAAGGG CCTCGTGATA
3001 CGCCTATTTT TATAGGTTAA TGTCA TGATA ATAATGGTTT CTTAGACGTC AGGTGGCACT
3061 TTTCGGGGAA ATGTGCGCGG AACCCCTATT TGTTTATTTT TCTAAATACA TTCAAATATG
3121 TATCCGCTCA TGAGACAATA ACCCTGATAA ATGCTTCAAT AATATTGAAA AAGGAAGAGT
3181 ATGAGTATTC AACATTTCCG TGTGCGCCCTT ATTCCCTTTT TTGCGGCATT TTGCTTCCT
3241 GTTTTGTCTC ACCCAGAAAC GCTGGTGAAA GTAAAAGATG CTGAAGATCA GTTGGGTGCA
3301 CGAGTGGGTT ACATCGAACT GGATCTCAAC AGCGGTAAGA TCCTTGAGAG TTTTCGCCCC
3361 GAAGAACGTT TTCCAATGAT GAGCACTTTT AAAGTTCTGC TATGTGGCGC GGTATTATCC
3421 CGTGTGACG CCGGGCAAGA GCAACTCGGT CGCCGCATAC ACTATTCTCA GAATGACTTG
3481 GTTGAGTACT CACCAGTCAC AGAAAAGCAT CTTACGGATG GCATGACAGT AAGAGAAATTA
3541 TGCACTGCTG CCATAACCAT GAGTGATAAC ACTGCGGCCA ACTTACTTCT GACAACGATC
3601 GGAGGACCGA AGGAGCTAAC CGCTTTTTTG CACAACATGG GGGATCATGT AACTCGCCTT
3661 GATCGTTGGG AACCGGAGCT GAATGAAGCC ATACCAAACG ACGAGCGTGA CACCACGATG
3721 CCTGCAGCAA TGGCAACAAC GTTGCAGCAA CTATTAAGTG GCGAACTACT TACTCTAGCT
3781 TCCCGGCAAC AATTAATAGA CTGGATGGAG GCGGATAAAG TTGCAGGACC ACTTCTGCGC
3841 TCGGCCCTTC CGGCTGGCTG GTTTATTGCT GATAAATCTG GAGCCGGTGA GCGTGGGTCT
3901 CGCGGTATCA TTGCAGCACT GGGGCCAGAT GGTAAGCCCT CCGTATCGT AGTTATCTAC
3961 ACGACGGGGA GTCAGGCAAC TATGGATGAA CGAAATAGAC AGATCGCTGA GATAGGTGCC
4021 TCACTGATTA AGCATTGGTA ACTGTCAGAC CAAGTTTACT CATATATACT TTAGATTGAT
4081 TTAAAACTTC ATTTTAAATT TAAAAGGATC TAGGTGAAGA TCCTTTTTGA TAATCTCATG
4141 ACCAAAATCC CTTAACGTGA GTTTTCGTTT CACTGAGCGT CAGACCCCGT AGAAAAGATC
4201 AAAGGATCTT CTTGAGATCC TTTTTTCTG CGCGTAATCT GCTGCTTGCA AACAAAAAAA
4261 CCACCGCTAC CAGCGGTGGT TTGTTTGCCG GATCAAGAGC TACCAACTCT TTTTCCGAAG
4321 GTAACCTGCT TCAGCAGAGC GCAGATACCA AATACTGTCC TTCTAGTGTA GCCGTAGTTA
4381 GGCCACCACT TCAAGAATC TGTAAGACCG CCTACATACC TCGCTCTGCT AATCCTGTTA
4441 CCAGTGGCTG CTGCCAGTGG CGATAAGTCG TGTCTTACCG GGTGGAATC AAGACGATAG
4501 TTACCGGATA AGGCGCAGCG GTCGGGCTGA ACGGGGGGTT CGTGACACA GCCCAGCTTG
4561 GAGCGAACGA CCTACACCGA ACTGAGATAC CTACAGCGTG AGCTATGAGA AAGCGCCACG
4621 CTTCGCCAAG GGAGAAAGGC GGACAGGTAT CCGGTAAGCG GCAGGGTCGG AACAGGAGAG
4681 CGCACGAGGG AGCTTCCAGG GGGAAACGCC TGGTATCTTT ATAGTCCTGT CGGGTTTCGC
4741 CACCTCTGAC TTGAGCGTCG ATTTTGTGTA TGCTCGTCAG GGGGGCGGAG CCTATGGAAA
4801 AACGCCAGCA ACGCGGCCTT TTTACGGTTT CTGGCCCTTT GCTGGCCTTT TGCTACATG
4861 TTCTTTCTG CGTTATCCCC TGATTCTGTG GATAACCGTA TTACCGCCTT TGAGTGAGCT
4921 GATACCGCTC GCCGCAGCCG AACGACCGAG CGCAGCGAGT CAGTGAGCGA GGAAGCGGAA
4981 GAGCGCCTGA TGCGGTATTT TCTCCTTACG CATCTGTGCG GTATTTTACA CCGCATATAT
5041 GGTGCACTCT CAGTACAATC TGCTCTGATG CCGCATAGTT AAGCCAGTAT ACACTCCGCT
5101 ATCGCTACGT GACTGGGTCA TGGCTGCGCC CCGACACCCG CCAACACCCG CTGACGCGCC
5161 CTGACGGGCT TGTCTGCTCC CGGCATCCGC TTACAGACAA GCTGTGACCG TCTCCGGGAG
5221 CTGCATGTGT CAGAGGTTTT CACCGTCATC ACCGAAACGC GCGAGGCAGC TGCGGTAAAG
5281 CTCATCAGCG TGGTCGTGAA GCGATTCACA GATGTCTGCC TGTTTATCCG CGTCCAGCTC
5341 GTTGAGTTTC TCCAGAAGCG TTAATGTCTG GCTTCTGATA AAGCGGGCCA TGTTAAGGGC
5401 GGTTTTTTCC TGTTTGGTCA CTGATGCCTC CGTGTAAGGG GGATTTCTGT TCATGGGGGT
5461 AATGATACCG ATGAAACGAG AGAGGATGCT CACGATACGG GTTACTGATG ATGAACATGC
5521 CCGGTTACTG GAACGTTGTG AGGGTAAACA ACTGGCGGTA TGGATGCGGC GGGACCAGAG
5581 AAAAATCACT CAGGGTCAAT GCCAGCGCTT CGTTAATACA GATGTAGGTG TTCCACAGGG
5641 TAGCCAGCAG CATCCTGCGA TGCAGATCCG GAACATAATG GTGCAGGGCG CTGACTTCCG
5701 CGTTTCCAGA CTTTACGAAA CACGGAAACC GAAGACCATT CATGTTGTTT CTCAGGTCCG
5761 AGACGTTTTG CAGCAGCAGT CGCTTACCGT TCGCTCGCGT ATCGGTGATT CATTCTGCTA
5821 ACCAGTAAGG CAACCCCGCC AGCCTAGCCG GGTCTCTAAC GACAGGAGCA CGATCATGCG
5881 CACCCGTGGC CAGGACCCAA CGCTGCCCGA GATGCGCCGC GTGCGGCTGC TGGAGATGGC
5941 GGACGCGATG GATATGTTCT GCCAAGGGTT GGTGTCGCA TTCACAGTTC TCCGCAAGAA
6001 TTGATTGGCT CCAATCTTGT GAGTGGTGAA TCCGTTAGCG AGGTGCCGCC GGCTTCCATT
6061 CAGGTCGAGG TGGCCCGGCT CCATGCACCG CGACGCAACG CGGGGAGGCA GACAAGGTAT
6121 AGGGCGGCGC CTACAATCCA TGCCAACCCG TTCCATGTGC TCGCCGAGGC GGCATAAATC-

```

FIGURE 44C

127/240

6181 GCCGTGACGA TCAGCGGTCC AGTGATCGAA GTTAGGCTGG TAAGAGCCGC GAGCGATCCT
6241 TGAAGCTGTC CCTGATGGTC GTCATCTACC TGCCTGGACA GCATGGCCTG CAACGCGGGC
6301 ATCCCGATGC CGCCGGAAGC GAGAAGAATC ATAATGGGGA AGGCCATCCA GCCTCGCGTC
6361 GCGAACGCCA GCAAGACGTA GCCCAGCGCG TCGGCCGCCA TGCCGGCGAT AATGGCCTGC
6421 TTCTCGCCGA AACGTTTGGT GCGGGGACCA GTGACGAAGG CTTGAGCGAG GCGGTGCAAG
6481 ATTCCGAATA CCGCAAGCGA CAGGCCGATC ATCGTCGCGC TCCAGCGAAA GCGGTCCTCG
6541 CCGAAAATGA CCCAGAGCGC TGCCGGCACC TGTCTACGA GTTGCATGAT AAAGAAGACA
6601 GTCATAAGTG CGGCGACGAT AGTCATGCCC CGCGCCCACC GGAAGGAGCT GACTGGGTTG
6661 AAGGCTCTCA AGGGCATCGG TCGATCGACG CTCTCCCTTA TCGGACTCCT GCATTAGGAA
6721 GCAGCCCAAGT AGTAGGTTGA GGCCGTTGAG CACCGCCGCC GCAAGGAATG GTGCATGCAA
6781 GGAGATGGCG CCCAACAGTC CCCCAGCCAC GGGGCGCTGCC ACCATACCCA CGCCGAAACA
6841 AGCGCTCATG AGCCCGAAGT GCGGAGCCCG ATCTTCCCCA TCGGTGATGT CGGCGATATA
6901 GCGGCCAGCA ACCGCACCTG TGGCGCCGGT GATGCCGGCC ACGATGCGTC CGGCGTAGAG
6961 G

FIGURE 44b

129/240

pDEST25 6652 bp

<u>Location (Base Nos.)</u>		<u>Gene Encoded</u>
844..720		attR1
953..1612		CmR
1732..1816		inactivated ccdA
1954..2259		ccdB
2300..2424		attR2
2432..2794		trx
1	CCGGAAGCGA GAAGAATCAT AATGGGGAAG GCCATCCAGC CTCGCGTCGC GAACGCCAGC	
61	AAGACGTAGC CCAGCGCGTC GGCCGCCATG CCGGCGATAA TGGCCTGCTT CTCGCCGAAA	
121	CGTTTGGTGG CGGGACCACT GACGAAGGCT TGAGCGAGGG CGTGCAAGAT TCCGAATACC	
181	GCAAGCGACA GGCCGATCAT CGTCGCGCTC CAGCGAAAGC GGTCCCTCGCC GAAAATGACC	
241	CAGAGCGCTG CCGGCACCTG TCCTACGAGT TGCATGATAA AGAAGACAGT CATAAGTGCG	
301	GCGACGATAG TCATGCCCCG CGCCACCCGG AAGGAGCTGA CTGGGTTGAA GGCTCTCAAG	
361	GGCATCGGTC GATCGACGCT CTCCCTTATG CGACTCCTGC ATTAGGAAGC AGCCCAGTAG	
421	TAGGTTGAGG CCGTTGAGCA CCGCCGCCGC AAGGAAATGGT GCATGCAAGG AGATGGCGCC	
481	CAACAGTCCC CCGGCCACGG GGCCTGCCAC CATACCCACG CCGAAACAAG CGCTCATGAG	
541	CCCGAAGTGG CGAGCCCGAT CTTCCCATC GGTGATGTCG GCGATATAGG CGCCAGCAAC	
601	CGCACCTGTG GCGCCGGTGA TGCCGGCCAC GATGCGTCCG GCGTAGAGGA TCGAGATCTC	
661	GATCCCGCGA AATTAATACG ACTCACTATA GGGAGACCAC AACGGTTTCC CTCTAGATCA	
721	CAAGTTTGTA CAAAAAAGCT GAACGAGAAA CGTAAAATGA TATAAATATC AATATATTAA	
781	ATTAGATTTT GCATAAAAAA CAGACTACAT AATACTGTAA AACACAACAT ATCCAGTCAC	
841	TATGGCGGCC GCATTAGGCA CCCAGGCTT TACACTTTAT GCTTCCGGCT CGTATAATGT	
901	GTGGATTTTG AGTTAGGATC CGCGGAGATT TTCAGGAGCT AAGGAAGCTA AAATGGAGAA	
961	AAAAATCACT GGATATACCA CCCTTGATAT ATCCCAATGG CATCGTAAAG AACATTTTGA	
1021	GGCATTTCAG TCAGTTGCTC AATGTACCTA TAACCAGACC GTTCAGCTGG ATATTACGGC	
1081	CTTTTTTAAAG ACCGTAAAGA AAAATAAGCA CAAGTTTTAT CCGGCCTTTA TTCACATTCT	
1141	TGCCCCCCTG ATGAATGCTC ATCCGGAATT CCGTATGGCA ATGAAAGACG GTGAGCTGGT	
1201	GATATGGGAT AGTGTTCACC CTTGTTACAC CGTTTTCCAT GAGCAAACTG AAACGTTTTT	
1261	ATCGCTCTGG AGTGAATACC ACGACGATTT CCGGCAGTTT CTACACATAT ATTTCGCAAGA	
1321	TGTGGCGTGT TACGGTGAAA ACCTGGCCTA TTTCCCTAAA GGGTTTTATTG AGAATATGTT	
1381	TTTCGTCTCA GCCAATCCCT GGGTGAGTTT CACCAGTTTT GATTTAAACG TGGCCAATAT	
1441	GGACAACCTC TTCGCCCCCG TTTTCACCAT GGGCAAATAT TATACGCAAG GCGACAAGGT	
1501	GCTGATGCCG CTGGCGATTC AGGTTTCATCA TGCCGTCTGT GATGGCTTCC ATGTCGGCAG	
1561	AATGCTTAAT GAATTACAAC AGTACTGCGA TGAGTGGCAG GCGCGGGCGT AAACGCGTGG	
1621	ATCCGGCTTA CTAAAAGCCA GATAACAGTA TGCCTATTG CCGCTGATT TTTGCGGTAT	
1681	AAGAATATAT ACTGATATGT ATACCCGAAG TATGTCAAAA AGAGGTGTGC TATGAAGCAG	
1741	CGTATTACAG TGACAGTTGA CAGCGACAGC TATCAGTTGC TCAAGGCATA TATGATGTCA	
1801	ATATCTCCGG TCTGGTAAGC ACAACCATGC AGAATGAAGC CCGTCGTCTG CGTGCCGAAC	
1861	GCTGGAAAGC GGAAAATCAG GAAGGGATGG CTGAGGTGCG CCGGTTTATT GAAATGAACG	
1921	GCTCTTTTGC TGACGAGAAC AGGGACTGGT GAAATGCAGT TTAAGGTTTA CACCTATAAA	
1981	AGAGAGAGCC GTTATCGTCT GTTTGTGGAT GTACAGAGTG ATATTATTGA CACGCCCGGG	
2041	CGACGGATGG TGATCCCCCT GGCCAGTGCA CGTCTGCTGT CAGATAAAGT CTCCCGTGAA	
2101	CTTTACCCGG TGGTGCAATAT CCGGGATGAA AGCTGGCGCA TGATGACCAC CGATATGGCC	
2161	AGTGTGCCGG TCTCCGTTAT CCGGGAAGAA GTGGCTGATC TCAGCCACCG CGAAAATGAC	
2221	ATCAAAAACG CCATTAACTT GATGTTCTGG GGAATATAAA TGTCAGGCTC CCTTATACAC	
2281	AGCCAGTCTG CAGGTCGACC ATAGTGACTG GATATGTTGT GTTTTACAGT ATTATGTAGT	
2341	CTGTTTTTTA TGCAAAATCT AATTTAATAT ATTGATATTT ATATCATTTT ACGTTTCTCG	
2401	TTTCAGCTTC TTGTACAAAG TGGTGATTAT GAGCGATAAA ATTATTCACC TGACTGACGA	
2461	CAGTTTTGAC ACGGATGTAC TCAAAGCGGA CGGGGCGATC CTCGTGATT TCTGGGCAGA	
2521	GTGGTGCGGT CCGTGCAAAA TGATCGCCCC GATTCTGGAT GAAATCGCTG ACGAATATCA	
2581	GGGCAAACTG ACCGTTGCAA AACTGAACAT CGATCAAAAC CCTGGCACTG CGCCGAAATA	
2641	TGGCATCCGT GGTATCCCGA CTCTGCTGCT GTTCAAAAAC GGTGAAGTGG CGGCAACCAA	
2701	AGTGGGTGCA CTGTCTAAAG GTCAGTTGAA AGAGTTTCTC GACGCTAACC TGGCCGGTTC	
2761	TGTTCTGGT GATGACGATG ACAAGGTACC CCGGGATCGA TCCGGCTGCT AACAAAGCCC	

FIGURE 45B

130/240

2821 GAAAGGAAGC TGAGTTGGCT GCTGCCACCG CTGAGCAATA ACTAGCATAA CCCCTTGGGG
2881 CCTCTAAACG GGTCTTGAGG GGTTTTTTGC TGAAAGGAGG AACTATATCC GGATATCCAC
2941 AGGACGGGTG TGGTCGCCAT GATCGCGTAG TCGATAGTGG CTCCAAGTAG CGAAGCGAGC
3001 AGGACTGGGC GGCGGCCAAA GCGGTCCGAC AGTGCTCCGA GAACGGGTGC GCATAGAAAT
3061 TGCATCAACG CATATAGCGC TAGCAGCACG CCATAGTGAC TGGCGATGCT GTCGGAATGG
3121 ACGATATCCC GCAAGAGGCC CGGCAGTACC GGCATAACCA AGCCTATGCC TACAGCATCC
3181 AGGGTGACGG TGCCGAGGAT GACGATGAGC GCATTGTTAG ATTTTCATACA CGGTGCCTGA
3241 CTGCGTTAGC AATTTAACTG TGATAAACTA CCGCATTAAA GCTTATCGAT GATAAGCTGT
3301 CAAACATGAG AATTCTTGAA GACGAAAGGG CCTCGTGATA CGCCTATTTT TATAGGTTAA
3361 TGTCATGATA ATAATGGTTT CTTAGACGTC AGGTGGCACT TTTCGGGGAA ATGTGCGCGG
3421 AACCCTTATT TGTTTATTTT TCTAAATACA TTCAAATATG TATCCGCTCA TGAGACAATA
3481 ACCCTGATAA ATGCTTCAAT AATATTGAAA AAGGAAGAGT ATGAGTATTC AACATTTCCG
3541 TGTCGCCCTT ATTCCCTTTT TTGCGGCATT TTGCCTTCCT GTTTTTGCTC ACCCAGAAAC
3601 GCTGGTGAAA GTAAAAGATG CTGAAGATCA GTTGGGTGCA CGAGTGGGTT ACATCGAACT
3661 GGTACTCAAC AGCGGTAAGA TCCTTGAGAG TTTTCGCCCC GAAGAACGTT TTCCAATGAT
3721 GAGCACTTTT AAAGTTCTGC TATGTGGCGC GGTATTATCC CGTGTGACG CCGGGCAAGA
3781 GCAACTCGGT CGCCGCATAC ACTATTCTCA GAATGACTTG GTTGAGTACT CACCAGTCAC
3841 AGAAAAGCAT CTTACGGATG GCATGACAGT AAGAGAATTA TGCAGTGCTG CCATAACCAT
3901 GAGTGATAAC ACTGCGGCCA ACTTACTTCT GACAACGATC GGAGGACCGA AGGAGCTAAC
3961 CGCTTTTTTG CACAACATGG GGGATCATGT AACTCGCCTT GATCGTTGGG AACCGGAGCT
4021 GAATGAAGCC ATACCAAACG ACGAGCGTGA CACCACGATG CCTGCAGCAA TGGCAACAAC
4081 GTTGCGCAAA CTATTAACCT GCGAACTACT TACTCTAGCT TCCCAGCAAC AATTAAATAGA
4141 CTGGATGGAG GCGGATAAAG TTGCAGGACC ACTTCTGCGC TCGGCCCTTC CGGCTGGCTG
4201 GTTTATTGCT GATAAATCTG GAGCCGGTGA GCGTGGGTCT CGCGGTATCA TTGCAGCACT
4261 GGGGCCAGAT GGTAAGCCCT CCCGTATCGT AGTTATCTAC ACGACGGGGA GTCAGGCAAC
4321 TATGGATGAA CGAAATAGAC AGATCGCTGA GATAGGTGCC TCACTGATTA AGCATTTGGTA
4381 ACTGTCAGAC CAAGTTTACT CATATATACT TTAGATTGAT TTAACCTTC ATTTTAAATT
4441 TAAAAGGATC TAGGTGAAGA TCCTTTTTGA TAATCTCATG ACCAAATCC CTTAACGTGA
4501 GTTTTCGTTT CACTGAGCGT CAGACCCCGT GCTGCTTGCA AACAAAAAAA CCACCGCTAC CAGCGGTGGT
4561 TTTTTCCTG CGCGTAATCT GCTGCTTGCA AACAAAAAAA CCACCGCTAC CAGCGGTGGT
4621 TTGTTTGCCG GATCAAGAGC TACCAACTCT TTTTCCGAAG GTAAGTGGCT TCAGCAGAGC
4681 GCAGATACCA AATACTGTCC TTCTAGTGTA GCCGTAGTTA GGCCACCACT TCAAGAACTC
4741 TGTAGCACCG CCTACATACC TCGCTCTGCT AATCCTGTTA CCAGTGGCTG CTGCCAGTGG
4801 CGATAAGTCG TGTCTTACCG GGTGAGCTC AAGACGATAG TTACCGGATA AGGCGCAGCG
4861 GTCGGGCTGA ACGGGGGGTT CGTGACACCA GCCCAGCTTG GAGCGAACGA CCTACACCGA
4921 ACTGAGATAC CTACAGCGTG AGCTATGAGA AAGCGCCACG CTTCCCGAAG GGAGAAAGGC
4981 GGACAGGTAT CCGGTAAGCG GCAGGGTCCG AACAGGAGAG CGCACGAGGG AGCTTCCAGG
5041 GGGAAACGCC TGGTATCTTT ATAGTCTGTG CGGGTTTCGC CACCTCTGAC TTGAGCGTCG
5101 ATTTTGTGA TGCTCGTCAG GGGGGCGGAG CCTATGGAAA AACGCCAGCA ACGCGGCCCTT
5161 TTTACGGTTC CTGGCCTTTT GCTGGCCTTT TGCTCATATG TTCTTTCCTG CGTTATCCCC
5221 TGATTCTGTG GATAACCGTA TTACCGCCTT TGAGTGAGCT GATACCGCTC GCCGCAGCCG
5281 AACGACCGAG CGCAGCGAGT CAGTGAGCGA GGAAGCGGAA GAGCGCCTGA TGCGGTATTT
5341 TCTCCTTACG CATCTGTGCG GTATTTTACA CCGCATATAT GGTGCACTCT CAGTACAATC
5401 TGCTCTGATG CCGCATAGTT AAGCCAGTAT ACACTCCGCT ATCGCTACGT GACTGGGTCA
5461 TGGCTGCGCC CCGACACCCG CCAACACCCG CTGACGCGCC CTGACGGGCT TGTCTGCTCC
5521 CGGCATCCGC TTACAGACAA GCTGTGACCG TCTCCGGGAG CTGCATGTGT CAGAGGTTTT
5581 CACCGTCATC ACCGAAACGC GCGAGGCAGC TGCGGTAAAG CTCATCAGCG TGGTCTGTGA
5641 GCGATTCACT GATGTCTGCC GTTTCATCCG CGTCCAGCTC GTTGAGTTTC TCCAGAAGCG
5701 TTAATGTCTG GCTTCTGATA AAGCGGGCCA TGTTAAGGGC GGTTTTTTCC TGTTTGGTCA
5761 CTGATGCCTC CGTGTAAGGG GGATTTCTGT TCATGGGGGT AATGATACCG ATGAAACGAG
5821 AGAGGATGCT CACGATACGG GTTACTGATG ATGAACATGC CCGGTACTG GAACGTTGTG
5881 AGGGTAAACA ACTGGCGGTA TGGATCGGCG GGGACCAGAG AAAAATCACT CAGGGTCAAT
5941 GCCAGCGCTT CGTTAATACA GATGTAGGTG TTCCACAGGG TAGCCAGCAG CATCCTGCGA
6001 TGCAGATCCG GAACATAATG GTGCAGGGCG CTGACTTCCG CGTTTCCAGA CTTTACGAAA
6061 CACGGAAACC GAAGACCATT CATGTTGTTG CTCAGGTGCG AGACGTTTTG CAGCAGCAGT
6121 CGCTTCACGT TCGCTCGCGT ATCGGTGATT CATTCTGCTA ACCAGTAAGG CAACCCCGCC
6181 AGCCTAGCCG GGTCTCAAC GACAGGAGCA CGATCATGCG CACCCGTGGC CAGGACCCAA
6241 CGCTGCCCGA GATGCGCCGC GTGCGGCTGC TGGAGATGGC GGACGCGATG GATATGTTCT

F6UR6 45C

131/240

6301 GCCAAGGGTT GGT TTGCGCA TTCACAGTTC TCCGCAAGAA TTGATTGGCT CCAATTCTTG
6361 GAGTGGTGAA TCCGTTAGCG AGGTGCCGCC GGCTTCCATT CAGGTCGAGG TGGCCCGGCT
6421 CCATGCACCG CGACGCAACG CGGGGAGGCA GACAAGGTAT AGGGCGGCGC CTACAATCCA
6481 TGCCAACCCG TTCCATGTGC TCGCCGAGGC GGCATAAATC GCCGTGACGA TCAGCGGTCC
6541 AGTGATCGAA GTTAGGCTGG TAAGAGCCGC GAGCGATCCT TGAAGCTGTC CCTGATGGTC
6601 GTCATCTACC TGCCTGGACA GCATGGCCTG CAACGCGGGC ATCCCGATGC CG

FIGURE 45D

132/240

FIGURE 46A

pDEST26 His6 Amino Fusion in pCMV Sport-neo Vector

```

600   ttg acg tca atg gga gtt tgt ttt ggc acc aaa atc aac ggg act ttc caa
      aac tgc agt tac cct caa aca aaa ccg tgg ttt tag ttg ccc tga aag gtc

651   aat gtc gta aca act ccg ccc oat tga cgc aaa tgg gcg gta ggc gtg tac
      tta cag cat tgt tga ggc ggg gta act gcg ttt acc cgc cat ccg cac atg

702   // ggt ggg agg tct ata taa goa gag etc gtt tag tga acc gtc aga tgc tct
      // cca ccc tcc aga tat att cgt ctc gag caa atc act tgg cag tct ago gga

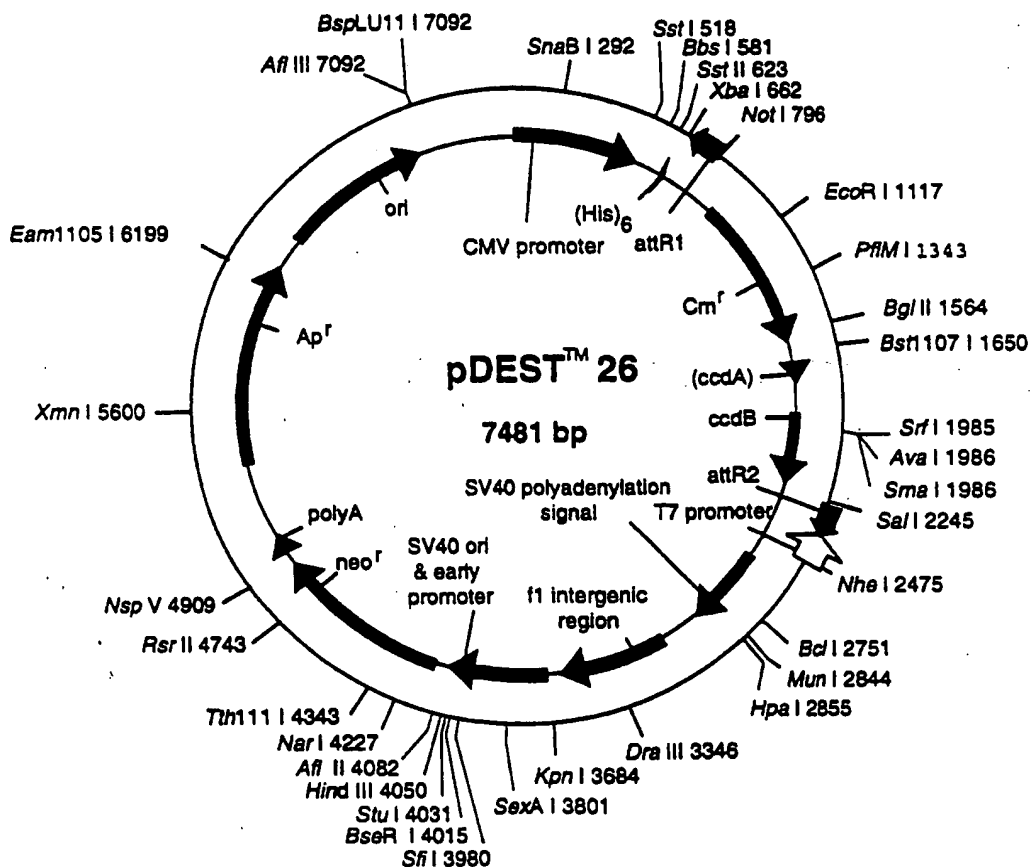
753   gga gac gcc atc cac gct gtt ttg acc tcc ata gaa gac acc ggg acc gat
      cct ctg cgg tag gtg cga caa aac tgg agg tat ctt ctg tgg ccc tgg cta

804   cca gcc tcc gga etc tag cct agg ccg cgg acc atg gcg tac tac cat cac
      ggt cgg agg cct gag atc gga tcc gcc gcc tgg tac cgc atg atg gta gtg

855   H H H H S R S   aca agt ttg tac aaa aaa gct gaa cga gaa
      gta gtg gta gtg aga tct agt   tgt tca aac atg ttt ttt cga att gct ctt
                                     Int Y

```

Handwritten annotations: "CMV Promoter" with an arrow pointing to the sequence between 651 and 702; "Start Transl" with an arrow pointing to the sequence between 753 and 804; "Int Y" with an arrow pointing to the sequence between 855 and 900.



pDEST26 7481 bp

<u>Location (Base Nos.)</u>	<u>Gene Encoded</u>
492..509	his6
619..519	attR1
752..1411	CmR
1531..1615	inactivated ccdA
1753..2058	ccdB
2099..2223	attR2
2409..2771	SV40 polyA
2966..3421	f1 intergenic region
3485..3903	SV40 promoter
3948..4742	neo
4806..4854	polyA
5265..6125	Apr
6274..6913	ori
7344..385	CMV promoter

```

1 GTAAACTGCC CACTTGGCAG TACATCAAGT GTATCATATG CCAAGTACGC CCCCTATTGA
61 CGTCAATGAC GGTAAATGGC CCGCCTGGCA TTATGCCAGC TACATGACCT TATGGGACTT
121 TCCTACTTGG CAGTACATCT ACGTATTAGT CATCGCTATT ACCATGGTGA TGC GGTTTGTG
181 GCAGTACATC AATGGGCGTG GATAGCGGTT TGA CTCACGG GGATTTCCTA GTCTCCACCC
241 CATTGACGTC AATGGGAGTT TGT TTTGGCA CCAAATCAA CGGGACTTTC CAAATGTCTG
301 TAACAATCCC GCCCCATTGA CGCAAATGGG CGGTAGGCGT GTACGGTGGG AGGTCTATAT
361 AAGCAGAGCT CGTTTAGTGA ACCGTCAGAT CGCCTGGAGA CGCCATCCAC GCTGTTTTGA
421 CCTCCATAGA AGACACCGGG ACCGATCCAG CCTCCGGACT CTAGCCTAGG CCGCGGACCA
481 TGGCGTACTA CCATCACCAT CACCATCACT CTAGATCAAC AAGTTTGTAC AAAAAAGCTG
541 AACGAGAAAC GTAAATGAT ATAAATATCA ATATATTAAA TTAGATTTTG CATAAAAAAC
601 AGACTACATA ATACTGTAAA ACACAACATA TCCAGTCACT ATGGCGGCCG CATTAGGCAC
661 CCCAGGCTTT ACAC TTTATG CTTCCGGCTC GTATAATGTG TGGATTTTGA GTTAGGATCC
721 GGCGAGATTT TCAGGAGCTA AGGAAGCTAA AATGGAGAAA AAAATCACTG GATATACCAC
781 CGTTGATATA TCCCAATGGC ATCGTAAAGA ACATTTTGAG GCATTTTCAGT CAGTTGCTCA
841 ATGTACCTAT AACCAGACCG TTCAGCTGGA TATTACGGCC TTTTAAAGA CCGTAAAGAA
901 AAATAAGCAC AAGTTTATC CGGCCTTTAT TCACATTCTT GCCCGCCTGA TGAATGCTCA
961 TCCGGAATTC CGTATGGCAA TGAAAGACGG TGAGCTGGTG ATATGGGATA GTGTTCAACC
1021 TTGTTACACC GTTTTCCATG AGCAAACCTGA AACGTTTTCA TCGCTCTGGA GTGAATACCA
1081 CGACGATTTT CGGCAGTTTC TACACATATA TTCGCAAGAT GTGGCGTGTT ACGGTGAAAA
1141 CCTGGCCTAT TTCCCTAAAG GGT TTTATTGA GAATATGTTT TTCGTCTCAG CCAATCCCTG
1201 GGTGAGTTTC ACCAGTTTTC ATTTAAACGT GGCCAATATG GACAACCTCT TCGCCCCCGT
1261 TTTCAACATG GGCAAATATT ATACGCAAGG CGACAAGGTG CTGATGCCGC TGGCGATTCA
1321 GGTTCATCAT GCGTCTGTG ATGGCTTCCA TGTCGGCAGA ATGCTTAATG AATTACAACA
1381 GTACTGCGAT GAGTGGCAGG GCGGGGCGTA AAGATCTGGA TCCGGCTTAC TAAAAGCCAG
1441 ATAACAGTAT GCGTATTTGC GCGCTGATTT TTGCGGTATA AGAATATATA CTGATATGTA
1501 TACCCGAAGT ATGTCAAAAA GAGGTGTGCT ATGAAGCAGC GTATTACAGT GACAGTTGAC
1561 AGCGACAGCT ATCAGTTGCT CAAGGCATAT ATGATGTCAA TATCTCCGGT CTGGTAAGCA
1621 CAACCATGCA GAATGAAGCC CGTCGTCTGC GTGCCGAACG CTGGAAAGCG GAAAATCAGG
1681 AAGGGATGGC TGAGGTCCGC CGGTTTATTG AAATGAACGG CTCTTTTGCT GACGAGAACA
1741 GGGACTGGTG AAATGCAGTT TAAGGTTTAC ACCTATAAAA GAGAGAGCCG TTATCGTCTG
1801 TTTGTGGATG TACAGAGTGA TATTATTGAC ACGCCCGGGC GACGGATGGT GATCCCCCTG
1861 GCCAGTGCAC GTCTGCTGTC AGATAAAGT TCCCGTGAAC TTTACCCGGT GGTGCATATC
1921 GGGGATGAAA GCTGGCGCAT GATGACCACC GATATGGCCA GTGTGCCGGT CTCCGTTATC
1981 GGGGAAGAAG TGGCTGATCT CAGCCACCGC GAAAATGACA TCAAAAACGC CATTAACTG
2041 ATGTTCTGGG GAATATAAAT GTCAGGCTCC CTTATACACA GCCAGTCTGC AGGTCGACCA
2101 TAGTGACTGG ATATGTTGTG TTTTACAGTA TTATGTAGTC TGTTTTTTAT GCAAAATCTA
2161 ATTTAATATA TTGATATTTA TATCATTTTA CGTTTCTCGT TCAGCTTTCT TGTACAAAGT
2221 GGTGATCGC GTGCATGCGA CGTCATAGCT CTCTCCCTAT AGTGAGTCGT ATTATAAGCT
2281 AGGCACTGGC CGTCGTTTTA CAACGTCGTG ACTGGGAAAA CTGCTAGCTT GGGATCTTTG -

```

Figure 46B

2341 TGAAGGAACC TTA CTTCTGT GGTGTGACAT AATTGGACAA ACTACCTACA GAGATTTAAA
2401 GCTCTAAGGT AAATATAAAA TTTTAAAGTG TATAATGTGT TAAACTAGCT GCATATGCTT
2461 GCTGCTTGAG AGTTTGTCTT ACTGAGTATG ATTTATGAAA ATATTATACA CAGGAGCTAG
2521 TGATTCTAAT TGTTTGTGTA TTTTAGATTG ACAGTCCCAA GGCTCATTTC AGGCCCCCTCA
2581 GTCCTCACAG TCTGTTTCATG ATCATAATCA GCCATACCAC ATTTGTAGAG GTTTTACTTG
2641 CTTTAAAAAA CCTCCACAC CTCCCCCTGA ACCTGAAACA TAAATGAAT GCAATTGTTG
2701 TTGTTAACTT GTTTATTGCA GCTTATAATG GTTACAAATA AAGCAATAGC ATCACAAATT
2761 TCACAAATAA AGCATTTTTT TCACTGCATT CTAGTTGTGG TTTGTCCAAA CTCATCAATG
2821 TATCTTATCA TGTCTGGATC GATCCTGCAT TAATGAATCG GCCAACGCGC GGGGAGAGGC
2881 GGTTTGCCTA TTGGCTGGCG TAATAGCGAA GAGGCCCGCA CCGATCGCCC TTCCCAACAG
2941 TTGCGCAGCC TGAATGGCGA ATGGGACGCG CCCTGTAGCG GCGCATTAG CGCGGCGGGT
3001 GTGGTGGTTA CGCGCAGCGT GACCGCTACA CTTGCCAGCG CCCTAGCGCC CGCTCCTTTC
3061 GCTTCTTCTT CTTCTTCTT CTCCACGTTT GCGCGCTTTC CCCGTCAAGC TCTAAATCGG
3121 GGGCTCCCTT TAGGGTTCCG ATTTAGTGCT TTACGGCACC TCGACCCCAA AAACTTGAT
3181 TAGGGTGATG GTTCACGTAG TGGGCCATCG CCCTGATAGA CGGTTTTTCG CCCTTTGACG
3241 TTGGAGTCCA CGTTCTTTAA TAGTGGACTC TTGTTCCAAA CTGGAACAAC ACTCAACCCT
3301 ATCTCGGTCT ATTCTTTTGA TTTATAAGGG ATTTTGCCGA TTTCCGCCCTA TTGGTTAAAA
3361 AATGAGCTGA TTTAACAAAT ATTTAACGCG AATTTTAAACA AAATATTAAC GTTTACAATT
3421 TCGCCTGATG CGGTATTTTC TCCTTACGCA TCTGTGCGGT ATTTTACACC GCATACGCGG
3481 ATCTGCGCAG CACCATGGCC TGAAATAACC TCTGAAAGAG GAACTTGGTT AGGTACCTTC
3541 TGAGGCGGAA AGAACCGCT GTGGAATGTG TGTCAGTTAG GGTGTGGAAA GTCCCCAGGC
3601 TCCCCAGCAG GCAGAAGTAT GCAAAGCATG CATCTCAATT AGTCAGCAAC CAGGTGTGGA
3661 AAGTCCCCAG GCTCCCCAGC AGGCAGAAAT ATGCAAAGCA TGCATCTCAA TTAGTCAGCA
3721 ACCATAGTCC CGCCCCTAAC TCCGCCATC CCGCCCCCTAA CTCCGCCAG TTCCGCCCAT
3781 TCCTCGCCCC ATGGCTGACT AATTTTTTTT ATTTATGCAG AGGCCGAGGC CGCCTCGGCC
3841 TCTGAGCTAT TCCAGAAGTA GTGAGGAGGC TTTTGTGGAG GCCTAGGCTT TTGCAAAAAG
3901 CTTGATTCTT CTGACACAAC AGTCTCGAAG TTAAGGCTAG AGCCACCATG ATTGAACAAG
3961 ATGGATTGCA CGCAGGTTCT CCGGCCGCTT GGGTGGAGAG GCTATTCGGC TATGACTGGG
4021 CACAACAGAC AATCGGCTGC TCTGATGCCG CCGTGTTCG GCTGTGAGCG CAGGGGCGCC
4081 CGGTTCTTTT TGTCAAGACC GACCTGTCCG GTGCCCTGAA TGAAGTGCAG GACGAGGCG
4141 CGCGGCTATC GTGGCTGGCC ACGACGGGCG TTCCTTGCGC AGCTGTGCTC GACGTTGTCA
4201 CTGAAGCGGG AAGGGACTGG CTGCTATTGG GCGAAGTGCC GGGGAGGAT CTCCTGTCTC
4261 CTCACCTTGC TCCTGCCGAG AAAGTATCCA TCATGGCTGA TGCAATGCGG CCGCTGCATA
4321 CGCTTGATCC GGCTACCTGC CCATTGACAC ACCAAGCGAA ACATCGCATC GAGCGAGCAC
4381 GTACTCGGAT GGAAGCCGGT CTTGTGATC AGGATGATCT GGACGAAGAG CATCAGGGGC
4441 TCGCGCGCAG CGAACTGTTT GCCAGGCTCA AGGCGCGCAT GCGGACGGC GAGGATCTCG
4501 TCGTGACCCA TGGCGATGCC TGCTTGCCGA ATATCATGGT GGAAAATGGC CGCTTTTCTG
4561 GATTTCATCGA CTGTGGCCGG CTGGGTGTGG CGGACCGCTA TCAGGACATA GCGTTGGCTA
4621 CCCGTGATAT TGCTGAAGAG CTTGGCGGCG AATGGGCTGA CCGCTTCTC GTGCTTTACG
4681 GTATCGCCGC TCCCGATTCC CAGCGCATCG CTTTCTATCG CTTTCTTGAC GAGTTCTTCT
4741 GAGCGGGACT CTGGGGTTCG AAATGACCGA CCAAGCGAGC CCAACCTGC CATCACGATG
4801 GCCGCAATAA AATATCTTTA TTTTCATTAC ATCTGTGTGT TGGTTTTTTG TGTGAATCGA
4861 TAGCGATAAG GATCCGCGTA TGGTGCACTC TCAGTACAAT CTGCTCTGAT GCCGATAGT
4921 TAAGCCAGCC CCGACACCCG CCAACACCCG CTGACGCGCC CTGACGGGCT TGTCTGCTCC
4981 CGGCATCCGC TTACAGACAA GCTGTGACCG TCTCCGGGAG CTGCATGTGT CAGAGGTTTT
5041 CACCGTCATC ACCGAAACGC GCGAGACGAA AGGGCCTCGT GATACGCCCTA TTTTATAGG
5101 TTAATGTCAT GATAATAATG GTTCTTTAGA CGTCAGGTGG CACTTTTTCG GGAATGTGC
5161 GCGGAACCCC TATTTGTTTA TTTTCTAAA TACATTCAAA TATGTATCCG CTCATGAGAC
5221 AATAACCTG ATAAATGCTT CAATAATATT GAAAAAGGAA GAGTATGAGT ATTCAACATT
5281 TCCGTGTGCG CTTTATTTCC TTTTGTGCG CATTTTGCCT TCCTGTTTTT GCTCACCAG
5341 AAACGCTGGT GAAAGTAAAA GATGCTGAAG ATCAGTTGGG TGCACGAGTG GGTACATCG
5401 AACTGGATCT CAACAGCGGT AAGATCCTTG AGAGTTTTTC CCCCAGAGAA CGTTTTCCAA
5461 TGATGAGCAC TTTTAAAGTT CTGCTATGTG GCGCGGTATT ATCCCGTATT GACGCGGGC
5521 AAGAGCAACT CGGTCGCGC ATACACTATT CTCAGAATGA CTTGGTTGAG TACTCACCAG
5581 TCACAGAAAA GCATCTTACG GATGCGATGA CAGTAAGAGA ATTATGAGT GCTGCCATAA
5641 CCATGAGTGA TAACACTGCG GCCAACTTAC TTCTGACAA GATCGGAGGA CCGAAGGAGC
5701 TAACCGCTTT TTTGCACAAC ATGGGGGATC ATGTAACCTG CTTGATCGT TGGGAACCGG
5761 AGCTGAATGA AGCCATACCA AACGACGAGC GTGACACCAC GATGCCTGTA GCAATGGCAA -

5821 CAACGTTGCG CAAACTATTA ACTGGCGAAC TACTTACTCT AGCTTCCCGG CAACAATTAA
5881 TAGACTGGAT GGAGGCGGAT AAAGTTGCAG GACCACTTCT GCGCTCGGCC CTTCCGGCTG
5941 GCTGGTTTAT TGCTGATAAA TCTGGAGCCG GTGAGCGTGG GTCTCGCGGT ATCATTGCAG
6001 CACTGGGGCC AGATGGTAAG CCTCCCCGTA TCGTAGTTAT CTACACGACG GGGAGTCAGG
6061 CAACTATGGA TGAACGAAAT AGACAGATCG CTGAGATAGG TGCCTCACTG ATTAAGCATT
6121 GGTAAC TGTC AGACCAAGTT TACTCATATA TACTTTAGAT TGATTAAAA CTTCATTTTT
6181 AATTTAAAAG GATCTAGGTG AAGATCCTTT TTGATAATCT CATGACCAAA ATCCCTTAAC
6241 GTGAGTTTTC GTTCCACTGA GCGTCAGACC CCGTAGAAAA GATCAAAGGA TCTTCTTGAG
6301 ATCCTTTTTT TCTGCGCGTA ATCTGCTGCT TGCAAACAAA AAAACCACCG CTACCAGCGG
6361 TGGTTTGTTC GCCGGATCAA GAGCTACCAA CTCTTTTTTCC GAAGGTAAC GGCTTCAGCA
6421 GAGCGCAGAT ACCAAATACT GTCCTTCTAG TGCTAGCCGTA GTTAGGCCAC CACTTCAAGA
6481 ACTCTGTAGC ACCGCCTACA TACCTCGCTC TGCTAATCCT GTTACCAGTG GCTGCTGCCA
6541 GTGGCGATAA GTCGTGTCTT ACCGGGTTGG ACTCAAGACG ATAGTTACCG GATAAGGCGC
6601 AGCGGTCGGG CTGAACGGGG GGTTCTGTGCA CACAGCCCAG CTTGGAGCGA ACGACCTACA
6661 CCGAACTGAG ATACCTACAG CGTGAGCATT GAGAAAGCGC CACGCTTCCC GAAGGGAGAA
6721 AGGCGGACAG GTATCCGGTA AGCGGCAGGG TCGGAACAGG AGAGCGCAGG AGGGAGCTTC
6781 CAGGGGGAAA CGCCTGGTAT CTTTATAGTC CTGTGCGGTT TCGCCACCTC TGACTTGAGC
6841 GTCGATTTTT GTGATGCTCG TCAGGGGGGC GGAGCCTATG GAAAAACGCC AGCAACGCGG
6901 CCTTTTTACG GTTCCTGGCC TTTTGCTGGC CTTTGCTCA CATGTTCTTT CCTGCGTTAT
6961 CCCCTGATTG TGTGGATAAC CGTATTACCG CCTTTGAGTG AGCTGATACC GCTCGCCGCA
7021 GCCGAACGAC CGAGCGCAGC GAGTCAGTGA GCGAGGAAGC GGAAGAGCGC CCAATACGCA
7081 AACCGCCTCT CCCCGCGCGT TGGCCGATTG ATTAATGCAG AGCTTGCAAT TCGCGCGTTT
7141 TTCAATATTA TTGAAGCATT TATCAGGGTT ATTGTCTCAT GAGCGGATAC ATATTTGAAT
7201 GTATTTAGAA AAATAAACAA ATAGGGGTTT CGCGCACATT TCCCCGAAAA GTGCCACCTG
7261 ACGTCTAAGA AACCATTATT ATCATGACAT TAACCTATAA AAATAGGCGT AGTACGAGGC
7321 CCTTTCACCTC ATTAGATGCA TGTGCTTACA TAACTTACGG TAAATGGCCC GCCTGGCTGA
7381 CCGCCCAACG ACCCCCGCCC ATTGACGTCA ATAATGACGT ATGTTCCCAT AGTAACGCCA
7441 ATAGGGACTT TCCATTGACG TCAATGGGTG GAGTATTTAC G

136/240
FIGURE 47A

pDEST 27 GST Amino Fusion in pCMV Sport-neo Vector

CMV Promoter

600 // nac ggt ggg agg tct ata taa gca gag ctc gtt tag tga acc gta aga tgg
ntg cca coc tcc aga taf att cgt ctc gag caa atc act tgg dag tct ago

651 cct gga gac gcc atc cac gct gtt ttg acv tcc ata gaa gac acc ggg acc
gga cct ctg cgg tag gtg cga caa aac tgg agg tat ctt atg tgg ccc tgg

702 gat cca gcc tcc gga clc tag cct agg cgg cgg acc atg gcc cct ata ata
cta ggt cgg agg cct gag atc gga tcc ggc gcc tgg tac cgg gga tat gat
Start Transl GST

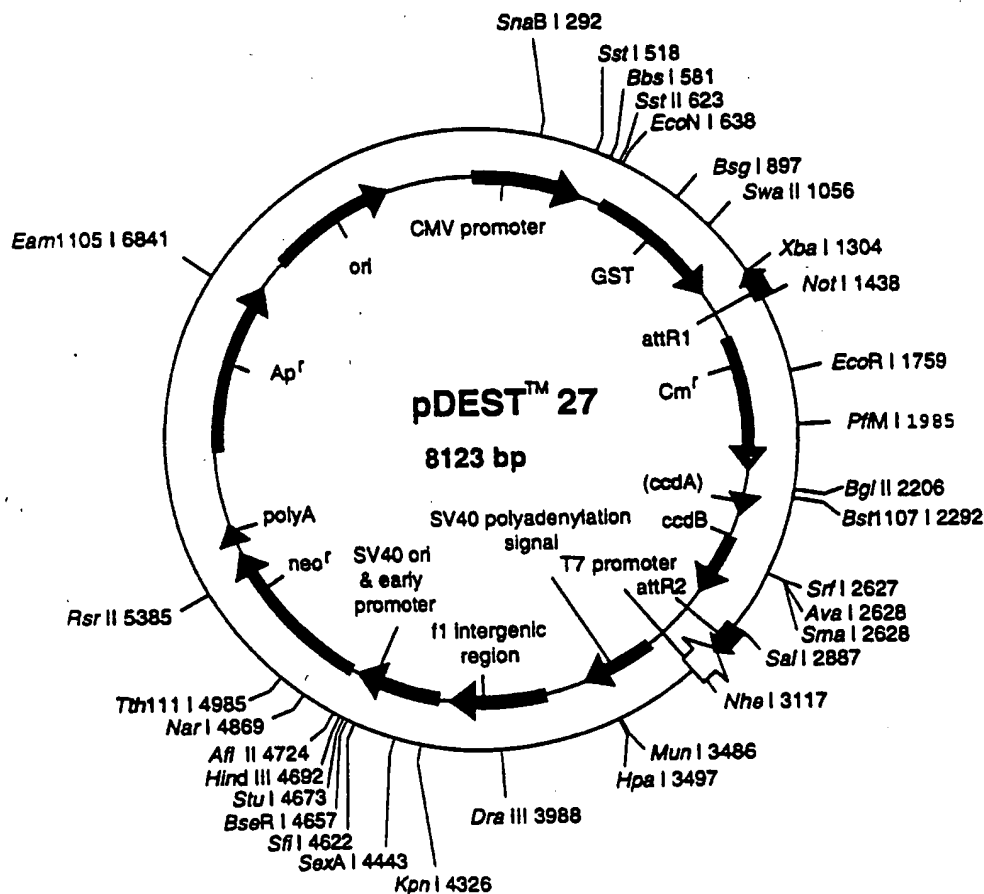
753 ggt tat tgg aaa att aag ggc ett gtg caa ccc act oga ott ctt ttg gaa
cca ata acc ttt taa ttc ccg gaa cac gtt ggg tga gct gaa gaa aac ctt

804 tat ctt gaa gaa aaa tat gaa gag cat ttg tat gag cgc gat gaa ggt gat
ata gaa ctt ctt ttt ata ctt ctc gta aac ata ctc gcg cta ctt cca cta

- GST Protein -

1365 // ttt ggt ggt ggc gac cat cct cca aaa tgg gat ctg gtt v P R S R
aaa cca cca cgg ctg gta gga ggt ttt agc cta gac oaa cgg cgt cct aga
S T S L Y K R A
1416 tca aca agt ttg tac aaa aaa gct gaa cga gaa aag
agt tgt tca aac atg ttg ttt cga ctt gct att tgc

Int attRI



137/240

pDEST27 8123 bp (rotated to position 7800)

<u>Location (Base Nos.)</u>	<u>Gene Encoded</u>
130..793	GST
803..927	attR1
1036..1695	CmR
1815..1899	inactivated ccdA
2037..2342	ccdB
2383..2507	attR2
2693..3055	SV40 polyA
3250..3705	f1 intergenic region
3769..4187	SV40 promoter
4232..5026	neo
5090..5138	polyA
5549..6409	Apr
6558..7197	ori
7628..27	CMV promoter

```

1 ATAAGCAGAG CTCGTTAGT GAACCGTCAG ATCGCCTGGA GACGCCATCC ACGTGTTTTT
61 GACCTCCATA GAAGACACCG GGACCGATCC AGCCTCCGGA CTCTAGCCTA GGCCGCGGAC
121 CATGGCCCCCT ATACTAGGTT ATTGGAATAA TAAGGGCCTT GTGCAACCCA CTCGACTTCT
181 TTTGGAATAT CTTGAAGAAA AATATGAAGA GCATTTGTAT GAGCGCGATG AAGGTGATAA
241 ATGGCGAAAC AAAAAGTTTG AATTGGGTTT GGAGTTTCCC AATCTTCTTT ATTATATTGA
301 TGGTGATGTT AAATTAACAC AGTCTATGGC CATCATACGT TATATAGCTG ACAAGCACAA
361 CATGTTGGGT GGTGTGCCAA AAGAGCGTGC AGAGATTTCA ATGCTTGAAG GAGCGGTTTT
421 GGATATTAGA TACGGTGTTT CGAGAATTGC ATATAGTAAA GACTTTGAAA CTCTCAAAGT
481 TGATTTTCTT AGCAAGCTAC CTGAAATGCT GAAAATGTTC GAAGATCGTT TATGTCATAA
541 AACATATTTA AATGGTGATC ATGTAACCCA TCCTGACTTC ATGTTGTATG ACGCTCTTGA
601 TGTGTTTTTA TACATGGACC CAATGTGCCT GGATGCGTTC CCAAAATTAG TTTGTTTTAA
661 AAAACGTATT GAAGCTATCC CACAAATTGA TAAGTACTTG AAATCCAGCA AGTATATAGC
721 ATGGCCTTTG CAGGGCTGGC AAGCCACGTT TGGTGGTGGC GACCATCCTC CAAAATCGGA
781 TCTGGTTCCG CGTTCTAGAT CAACAAGTTT GTACAAAAAA GCTGAACGAG AAACGTAAAA
841 TGATATAAAT ATCAATATAT TAAATTAGAT TTTGCATAAA AAACAGACTA CATAACTCTG
901 TAAACACAA CATATCCAGT CACTATGGCG GCCGCATTAG GCACCCGAGG CTTTACACTT
961 TATGCTTCCG GCTCGTATAA TGTGTGGATT TTGAGTTAGG ATCCGGCGAG ATTTTCAGGA
1021 GCTAAGGAAG CTAAAATGGA GAAAAAATC ACTGGATATA CCACCGTTGA TATATCCCAA
1081 TGGCATCGTA AAGAACATTT TGAGGCATTT CAGTCAGTTG CTCAATGTAC CTATAACCAG
1141 ACCGTTTCAG TGGATATTAC GGCCTTTTTA AAGACCGTAA AGAAAAATAA GCACAAGTTT
1201 TATCCGGCCT TTATTCACAT TCTTGCCCGC CTGATGAATG CTCATCCGGA ATTCCGTATG
1261 GCAATGAAAG ACGGTGAGCT GGTGATATGG GATAGTGTTC ACCCTTGTTA CACCGTTTTT
1321 CATGAGCAAA CTGAAACGTT TTCATCGCTC TGGAGTGAAT ACCACGACGA TTTCCGGCAG
1381 TTTCTACACA TATATTCGCA AGATGTGGCG TGTTACGGTG AAAACCTGGC CTATTTCCCT
1441 AAAGGGTTTA TTGAGAATAT GTTTTTCGTC TCAGCCAATC CCTGGGTGAG TTTCAACAGT
1501 TTTGATTTAA ACGTGGCCAA TATGGACAAC TTCTTCGCCC CCGTTTTTAC CATGGGCAAA
1561 TATTATACGC AAGGCGACAA GGTGCTGATG CCGCTGGCGA TTCAGGTTCA TCATGCCGTC
1621 TGTGATGGCT TCCATGTCCG CAGAATGCTT AATGAATTAC AACAGTACTG CGATGAGTGG
1681 CAGGGCGGGG CGTAAAGATC TGGATCCGGC TTAATAAAG CCAGATAACA GTATGCGTAT
1741 TTGCGCGCTG ATTTTTGCGG TATAAGAATA TATACTGATA TGTATACCCG AAGTATGTCA
1801 AAAAGAGGTG TGCTATGAAG CAGCGTATTA CAGTGACAGT TGACAGCGAG AGCTATCAGT
1861 TGCTCAAGGC ATATATGATG TCAATATCTC CGGTCTGGTA AGCACAACCA TGCAGAATGA
1921 AGCCCGTCGT CTGCGTGCCG AACGCTGGAA AGCGGAAAAT CAGGAAGGGA TGGCTGAGGT
1981 CGCCCGGTTT ATTGAAATGA ACGGCTCTTT TGCTGACGAG AACAGGGACT GGTGAAATGC
2041 AGTTTAAGGT TTACACCTAT AAAAGAGAGA GCCGTTATCG TCTGTTTGTG GATGTACAGA
2101 GTGATATTAT TGACACGCCC GGGCGACGGA TGGTGATCCC CCTGGCCAGT GCACGTCTGC
2161 TGTCAGATAA AGTCTCCCGT GAACTTTACC CGGTGGTGCA TATCGGGGAT GAAAGCTGGC
2221 GCATGATGAC CACCGATATG GCCAGTGTGC CGGTCTCCGT TATCGGGGAA GAAGTGGCTG
2281 ATCTCAGCCA CCGCGAAAAT GACATCAAAA ACGCCATTAA CCTGATGTTT TGGGGAATAT-

```

FIGURE 47B

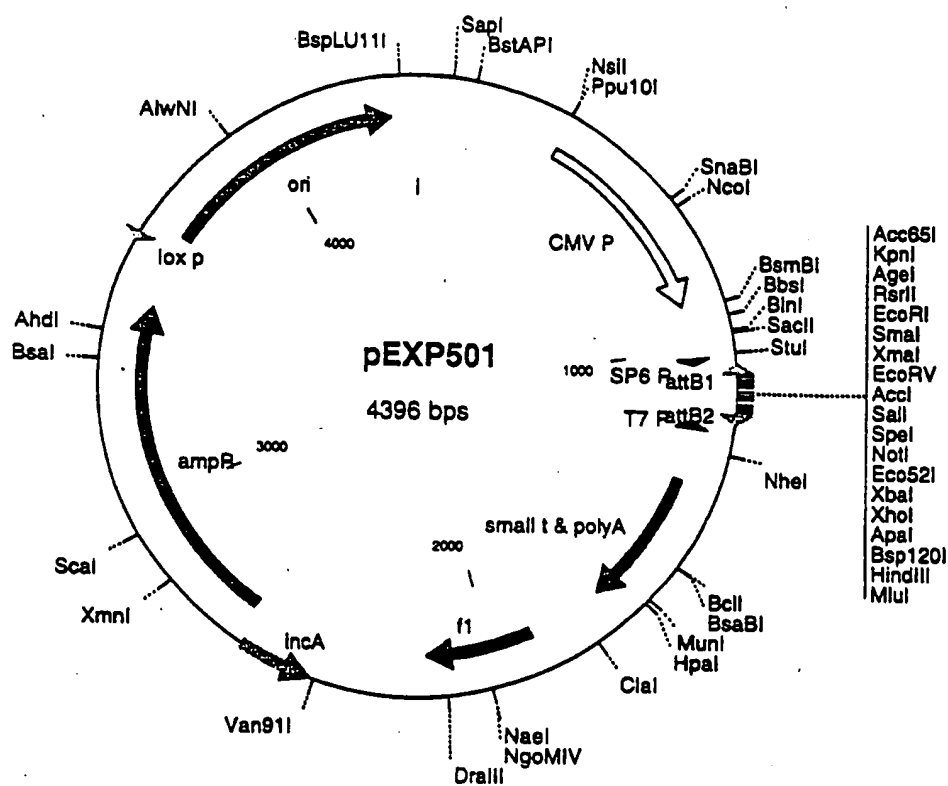
138/240

2341 AAATGTCAGG CTCCTTATA CACAGCCAGT CTGCAGGTCG ACCATAGTGA CTGGATATGT
2401 TGTGTTTTAC AGTATTATGT AGTCTGTTTT TTATGCAAAA TCTAATTTAA TATATTGATA
2461 TTTATATCAT TTTACGTTTC TCGTTCAGCT TTCTTGTTACA AAGTGTTTGA TCGCGTGCGT
2521 GCGACGTCAT AGCTCTCTCC CTATAGTGAG TCGTATTATA AGCTAGGCAC TGGCCGTCGT
2581 TTTACAACGT CGTGACTGGG AAAACTGCTA GCTTGGGATC TTTGTGAAGG AACCTTACTT
2641 CTGTGGTGTG ACATAATTGG ACAAACCTACC TACAGAGATT TAAAGCTCTA AGGTAAATAT
2701 AAAATTTTTA AGTGATAAT GTGTTAAACT AGCTGCATAT GCTTGCTGCT TGAGAGTTTT
2761 GCTTACTGAG TATGATTTAT GAAAATATTA TACACAGGAG CTAGTGATTC TAATTGTTTG
2821 TGTATTTTAG ATTACAGTC CCAAGGCTCA TTTCAGGCCC CTCAGTCCTC ACAGTCTGTT
2881 CATGATCATA ATCAGCCATA CCACATTTGT AGAGGTTTTA CTTGCTTTAA AAAACCTCCC
2941 ACACCTCCCC CTGAACCTGA AACATAAAAT GAATGCAATT GTTGTGTTA ACTTGTTTAT
3001 TGCAGCTTAT AATGGTTACA AATAAAGCAA TAGCATCACA AATTTACAA ATAAAGCATT
3061 TTTTTCACGT CATCTAGTT GTGGTTTGTG CAAACTCATC AATGTATCTT ATCATGTCTG
3121 GATCGATCCT GCATTAATGA ATCGGCCAAC GCGCGGGGAG AGGCGGTTTG CGTATTGGCT
3181 GCGGTAATAG CGAAGAGGCC CGCACCGATC GCCCTTCCCA ACAGTTGCGC AGCCTGAATG
3241 GCGAATGGGA CGCGCCCTGT AGCGGCGCAT TAAGCGCGGC GGGTGTGGTG GTTACGCGCA
3301 GCGTGACCGC TACACTTGCC AGCGCCCTAG CGCCCGCTCC TTTCTGCTTC TCCCTTCCT
3361 TTCTCGCCAC GTTCGCGCGC TTTCCCGGTC AAGCTCTAAA TCGGGGGCTC CCTTTAGGGT
3421 TCCGATTTAG TGCTTTACGG CACCTCGACC CCAAAAACT TGATTAGGGT GATGCGTTCAC
3481 GTAGTGGGCC ATCGCCCTGA TAGACGGTTT TTCGCCCTTT GACGTGAGG TCCAGTTCT
3541 TTAATAGTGG ACTCTTGTT CAAACTGGAA CAACACTCAA CCCTATCTCG GTCTATTCTT
3601 TTGATTTATA AGGGATTTTG CCGATTTTCG CCTATTGGTT AAAAAATGAG CTGATTTAAC
3661 AAATATTTAA CGCGAATTTT AACAAAATAT TAACGTTTAC AATTTGCGCT GATGCGGTAT
3721 TTTCTCCTTA CGCATCTGTG CCGTATTTCA CACCGCATAC GCGGATCTGC GCAGCACCAT
3781 GGCCTGAAAT AACCTCTGAA AGAGGAACTT GGTTAGGTAC CTTCTGAGGC GGAAAGAACC
3841 AGCTGTGGAA TGTGTGTCAG TTAGGGTGTG GAAAGTCCCC AGGCTCCCCA GCAGGCAGAA
3901 GTATGCAAAG CATGCATCTC AATTAGTCAG CAACCAGGTG TGGAAAGTCC CCAGCTCCCC
3961 CAGCAGGCAG AAGTATGCAA AGCATGCATC TCAATTAGTC AGCAACCATA GTCCCGCCCC
4021 TAACTCCGCC CATCCCGCCC CTAACCTCCG CCAGTTCCGC CCATTCTCCG CCCCATGGCT
4081 GACTAATTTT TTTTATTTAT GCAGAGGCCG AGGCCGCCCTC GGCCTCTGAG CTATTCCAGA
4141 AGTAGTGAGG AGGCTTTTTT GGAGGCCTAG GCTTTTGCAA AAAGCTTGAT TCTTCTGACA
4201 CAACAGTCTC GAACTTAAGG CTAGAGCCAC CATGATTGAA CAAGATGGAT TGCACGCAGG
4261 TTCTCCGGCC GCTTGGGTGG AGAGGCTATT CGGCTATGAC TGGGCACAAC AGACAATCGG
4321 CTGCTCTGAT GCCGCCGTGT TCCGGCTGTC AGCGCAGGGG CGCCCGGTTT TTTTGTCAA
4381 GACCGACCTG TCCGGTGCCC TGAATGAACT GCAGGACGAG GCAGCGCGGC TATCGTGGCT
4441 GGCCACGACG GGCCTTCCTT GCGCAGCTGT GCTCGACGTT GTCACTGAAG CGGGAAGGGA
4501 CTGGCTGCTA TTGGGCGAAG TGCCGGGGCA GGATCTCCTG TCATCTCACC TTGCTCCTGC
4561 CGAGAAAGTA TCCATCATGG CTGATGCAAT GCGGCGGCTG CATACTGCTG ATCCGGCTAC
4621 CTGCCCATTG GACCACCAAG CGAAACATCG CATCGAGCGA GCACGTACTC GGATGGAAGC
4681 CGGTCTTGTC GATCAGGATG ATCTGGACGA AGAGCATCAG GGGCTCGCGC CAGCCGAAC
4741 GTTCGCCAGG CTCAAGGCGC GCATGCCCGA CGGCGAGGAT CTCGTCGTGA CCCATGGCGA
4801 TGCCTGCTTG CCGAATATCA TGGTGGAAAA TGGCCGCTTT TCTGGATTCA TCGACTGTGG
4861 CCGCTGTGGT GTGGCGGACC GCTATCAGGA CATAGCGTTG GCTACCCGTG ATATTGCTGA
4921 AGAGCTTGCG GCGGAATGGG CTGACCGCTT CCTCGTGCTT TACGGTATCG CCGCTCCCGA
4981 TTCGCAGCGC ATCGCCTTCT ATCGCCTTCT TGACGAGTTC TTCTGAGCGG GACTCTGGGG
5041 TTCGAAATGA CCGACCAAGC GACGCCCAAC CTGCCATCAC GATGGCCGCA ATAAATATC
5101 TTTATTTTCA TTACATCTGT GTGTTGGTTT TTTGTGTGAA TCGATAGCGA TAAGGATCCG
5161 CGTATGGTGC ACTCTCAGTA CAATCTGCTC TGATGCCGCA TAGTTAAGCC AGCCCCGACA
5221 CCCGCCAACA CCCGCTGACG CGCCCTGACG GGCTTGCTCT CTCCCGCAT CCGCTTACAG
5281 ACAAGCTGTG ACCGTCTCCG GGAGCTGCAT GTGTCAGAGG TTTTACCCTG CATCACCGAA
5341 ACGCGCGAGA CGAAAGGGCC TCGTGATACG CCTATTTTTA TAGGTTAATG TCATGATAAT
5401 AATGGTTTCT TAGACGTCAG GTGGCACTTT TCGGGGAAAT GTGCGCGGAA CCCCTATTTG
5461 TTTATTTTTT TAAATACATT CAAATATGTA TCCGCTCATG AGACAATAAC CCTGATAAAT
5521 GCTTCAATAA TATTGAAAAA GGAAGAGTAT GAGTATTCAA CATTTCCGTG TCGCCCTTAT
5581 TCCCTTTTTT GCGGCATTTT GCCTTCCTGT TTTGCTCAC CCAGAAACGC TGGTGAAAGT
5641 AAAAGATGCT GAAGATCAGT TGGGTGCACG AGTGGGTAC ATCGAACTGG ATCTCAACAG
5701 CGGTAAGATC CTTGAGAGTT TTCGCCCCGA AGAACGTTTT CCAATGATGA GCACCTTTAA
5761 AGTTCTGCTA TGTGGCGCGG TATTATCCCG TATTGACGCC GGGCAAGAGC AACTCGGTCTG

Figure 47C

5821 CCGCATACAC TATTCTCAGA ATGACTTGGT TGAGTACTCA CCAGTCACAG AAAAGCATCT
5881 TACGGATGGC ATGACAGTAA GAGAATTATG CAGTGCTGCC ATAACCATGA GTGATAACAC
5941 TGCGGCCAAC TTACTTCTGA CAACGATCGG AGGACCGAAG GAGCTAACCG CTTTTTTGCA
6001 CAACATGGGG GATCATGTAA CTCGCCTTGA TCGTTGGGAA CCGGAGCTGA ATGAAGCCAT
6061 ACCAAACGAC GAGCGTGACA CCACGATGCC TGTAGCAATG GCAACAACGT TGCGCAAACT
6121 ATTAAC TGGC GAAC TACTTA CTCTAGCTTC CCGGCAACAA TTAATAGACT GGATGGAGGC
6181 GGATAAAGTT GCAGGACCAC TTCTGCGCTC GGCCCTTCCG GCTGGCTGGT TTATTGCTGA
6241 TAAATCTGGA GCCGGTGAGC GTGGGTCTCG CGGTATCATT GCAGCACTGG GGCAGATGG
6301 TAAGCCCTCC CGTATCGTAG TTATCTACAC GACGGGGAGT CAGGCAACTA TGGATGAACG
6361 AAATAGACAG ATCGCTGAGA TAGGTGCCTC ACTGATTAAG CATTGGTAAC TGTGAGACCA
6421 AGTTTACTCA TATATACTTT AGATTGATTT AAAACTTCAT TTTTAATTTA AAAGGATCTA
6481 GGTGAAGATG CTTTTTGATA ATCTCATGAC CAAAATCCCT TAACGTGAGT TTTGCTTCCA
6541 CTGAGCGTCA GACCCCGTAG AAAAGATCAA AGGATCTTCT TGAGATCCTT TTTTCTGCG
6601 CGTAATCTGC TGCTTGCAA CAAAAAACC ACCGCTACCA GCGGTGGTTT GTTTGCCGGA
6661 TCAAGAGCTA CCAACTCTTT TTCCGAAGGT AACTGGCTTC AGCAGAGCGC AGATACCAAA
6721 TACTGTCTTT CTAGTGTAGC CGTAGTTAGG CCACCACCTC AAGAACTCTG TAGCACC GCC
6781 TACATACCTC GCTCTGCTAA TCCTGTTACC AGTGGCTGCT GCCAGTGGCG ATAAGTCGTG
6841 TCTTACCGGG TTGGACTCAA GACGATAGTT ACCGGATAAG GCGCAGCGGT CCGGCTGAAC
6901 GGGGGGTTTC TGACACAGC CCAGCTTGGA GCGAACGACC TACCCGAAC TGAGATACCT
6961 ACAGCGTGAG CATTGAGAAA GCGCCACGCT TCCCGAAGGG AGAAAGGCGG ACAGGTATCC
7021 GGTAAGCGGC AGGGTCGGAA CAGGAGAGCG CACGAGGGAG CTTCCAGGGG GAAACGCCTG
7081 GTATCTTTAT AGTCCTGTCG GGTTTCGCCA CCTCTGACTT GAGCGTCGAT TTTTGTGATG
7141 CTCGTGAGGG GGGCGGAGCC TATGGAAAA CGCCAGCAAC GCGGCCTTTT TACGGTCTCT
7201 GGCCTTTTGC TGGCCTTTTG CTCACATGTT CTTTCCTGCG TTATCCCTG ATTCTGTGGA
7261 TAACCGTATT ACCGCCCTTG AGTGAGCTGA TACCGCTCGC CGCAGCCGAA CGACCGAGCG
7321 CAGCGAGTCA GTGAGCGAGG AAGCGGAAGA GCGCCCAATA CGCAAACCGC CTCTCCCCGC
7381 GCGTTGGCCG ATTCATTAAT GCAGAGCTTG CAATTGCGCG GTTTTTCAAT ATTATTGAAG
7441 CATTTATCAG GGTTATTGTC TCATGAGCGG ATACATATTT GAATGTATTT AGAAAAATAA
7501 ACAAATAGGG GTTCCGCGCA CATTTCCCCG AAAAGTGCCA CCTGACGTCT AAGAAACCAT
7561 TATTATCATG ACATTAACTT ATAAAAATAG GCGTAGTACG AGGCCCTTTC ACTCATTAGA
7621 TGCATGTCGT TACATAACTT ACGGTAAATG GCGCGCCTGG CTGACCGCCC AACGACCCCC
7681 GCCCATTGAC GTCAATAATG ACGTATGTTT CCATAGTAAC GCCAATAGGG ACTTTCCATT
7741 GACGTCAATG GGTGGAGTAT TTACGGTAAA CTGCCCCTT GGCAGTACAT CAAGTGTATC
7801 ATATGCCAAG TACGCCCCCT ATTGACGTCA ATGACGGTAA ATGGCCCGCC TGGCATTATG
7861 CCCAGTACAT GACCTTATGG GACTTTCCTA CTTGGCAGTA CATCTACGTA TTAGTCATCG
7921 CTATTACCAT GGTGATGCGG TTTTGGCAGT ACATCAATGG GCGTGGATAG CGGTTTGAAT
7981 CACGGGGATT TCCAAGTCTC CACCCCATG ACGTCAATGG GAGTTTGT TT TGGCACCAA
8041 ATCAACGGGA CTTTCCAAAA TGTCGTAACA ACTCCGCCCC ATTGACGCAA ATGGGCGGTA
8101 GCGGTGTACG GTGGGAGGTC TAT

140/240

Figure 4B A: pEXP501: pCMV.SPORT 6 host for attB Libraries

142/240

pEXP501 4396 bp

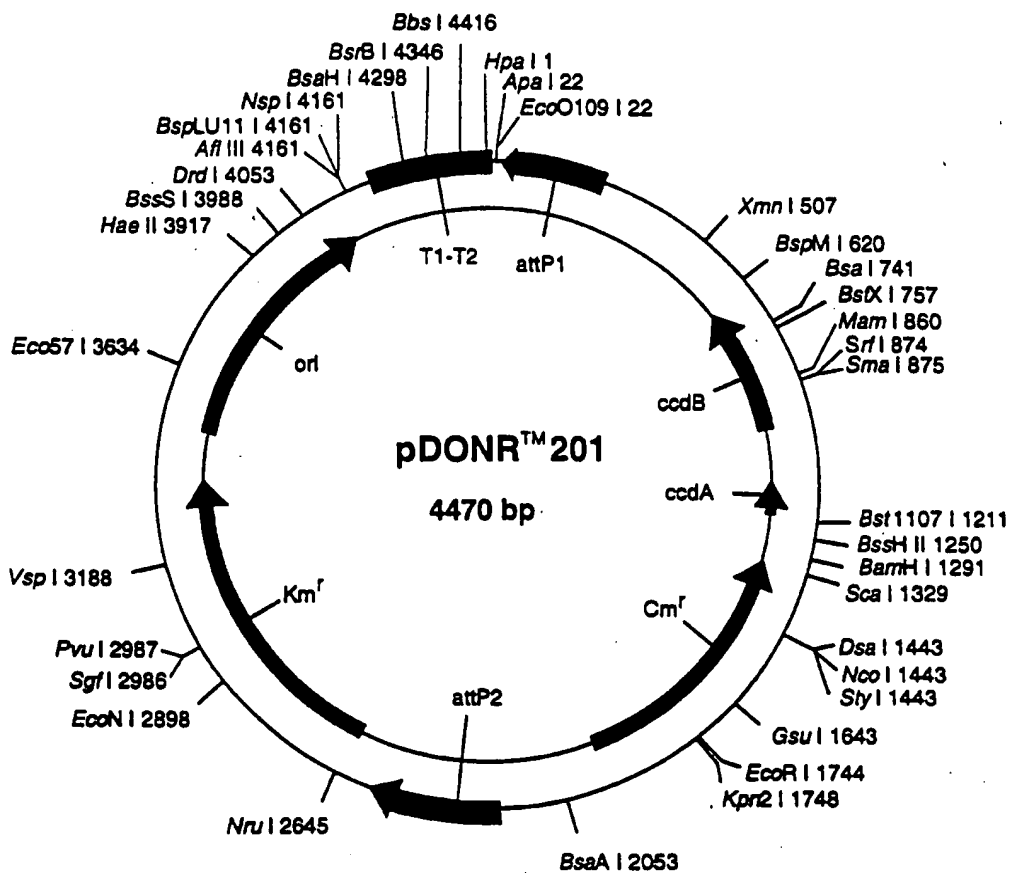
1 CCATTGCGCA TTCAGGCTGC GCAACTGTTG GGAAGGGCGA TCGGTGCGGG CCTCTTCGCT
61 ATTACGCCAG CCAATACGCA AACCGCCTCT CCCCGCGCGT TGGCCGATTC ATTAATGCAG
121 GATCGATCCA GACATGATAA GATACATTGA TGAGTTTGGG CAAACCACAA CTAGAATGCA
181 GTGAAAAAAA TGCTTTATTT GTGAAATTTG TGATGCTATT GCTTTATTTG TAACCATTAT
241 AAGCTGCAAT AAACAAGTTA ACAACAACAA TTGCATTCAT TTTATGTTTC AGGTTTCAGGG
301 GGAGGTGTGG GAGGTTTTTT AAAGCAAGTA AACCTCTAC AAATGTGGTA TGGCTGATTA
361 TGATCATGAA CAGACTGTGA GGAAGGAGG GCCTGAAATG AGCCTTGGGA CTGTGAATCT
421 AAAATACACA AACAATTAGA ATCAGTAGCT CCTGTGTATA ATATTTTCAT AAATCATACT
481 CAGTAAGCAA AACTCTCAAG CAGCAAGCAT ATGCAGCTAG TTAAACACAT TATACACTTA
541 AAAATTTTAT ATTTACCTTA GAGCTTTAAA TCTCTGTAGG TAGTTTGTCC AATTATGTCA
601 CACCACAGAA GTAAGGTTCC TTCACAAGA TCCCAAGCTA GCAGTTTTCC CAGTCACGAC
661 GTTGTAAAAA GACGGCCAGT GCCTAGCTTA TAATACGACT CACTATAGGG ACCACTTTGT
721 ACAAGAAAGC TGGGTACGCG TAAGCTTGGG CCCCTCGAGG GATCCTCTAG AGCGGCCGCC
781 GACTAGTGAG CTCGTGACG ATATCCCGGG AATTCCGGAC CGGTACCAGC CTGCTTTTTT
841 GTACAAACTT GTTCTATAGT GTCACCTAAA TAGGCCTAAT GGTCATAGCT GTTTCCTGTG
901 TGAAATTGTT ATCCGCTCCG CGGCCTAGGC TAGAGTCCGG AGGCTGGATC GGTCCCGGTG
961 TCTTCTATGG AGGTCAAAAC AGCGTGGATG GCGTCTCCAG GCGATCTGAC GGTTCATAA
1021 ACGAGCTCTG CTTATATAGA CCTCCCACCG TACACGCCTA CCGCCCATTT GCGTCAATGG
1081 GGC GGAGTTG TTACGACATT TTGGAAGTTC CCGTTGATTT TGGTGCCAAA ACAAACTCCC
1141 ATTGACGTCA ATGGGGTGGG GACTTGGAAA TCCCCGTGAG TCAAACCGCT ATCCACGCCC
1201 ATTGATGTAC TGCCAAAACC GCATCACCAT GGTAAATAGC ATGACTAATA CGTAGATGTA
1261 CTGCCAAGTA GGAAAGTCCC ATAAGGTCAT GTACTGGGCA TAATGCCAGG CGGGCCATTT
1321 ACCGTCAATTG ACGTCAATAG GGGGCGTACT TGGCATATGA TACACTTGAT GTACTGCCAA
1381 GTGGGCAGTT TACCGTAAAT ACTCCACCCA TTGACGTCAA TGGAAAGTCC CTATTGGCGT
1441 TACTATGGGA ACATACGTCA TTATTGACGT CAATGGGCGG GGGTCGTTGG GCGGTCAGCC
1501 AGGCGGGCCA TTTACCGTAA GTTATGTAA GACATGCATC TAATGAGTGA AAGGGCCTCG
1561 TACTACGCCT ATTTTATAG GTTAATGTCA TGATAATAAT GGTTCCTTAG ACGTCAGGTG
1621 GCACTTTTTC GGGAAATGTG CGCGGAACCC CTATTTGTTT ATTTTCTAA ATACATTCAA
1681 ATATGTATCC GCTCATGAGA CAATAACCCT GATAAATGCT TCAATAATAT TGAAAAACGC
1741 GCGAATTGCA AGCTCTGCAT TAATGAATCG GCCAACGCGC GGGGAGAGGC GGTTCGCGTA
1801 TTGGGCGCTC TTCCGCTTCC TCCTCACTG ACTCGCTGCG CTCGGTCGTT CGGCTGCGGC
1861 GAGCGGTATC AGCTCACTCA AAGGCGGTAA TACGGTTATC CACAGAATCA GGGGTAACG
1921 CAGGAAAGAA CATGTGAGCA AAAGGCCAGC AAAAGGCCAG GAACCGTAAA AAGGCCGCGT
1981 TGCTGGCGTT TTTCCATAGG CTCGCCCCC CTGACGAGCA TCACAAAAAT CGACGCTCAA
2041 GTCAGAGGTG GCGAAACCCG ACAGGACTAT AAAGATACCA GGCCTTTCCC CCTGGAAGCT
2101 CCTCGTGCCT CTCTCCTGTT CCGACCCTGC CGCTTACCGG ATACCTGTCC GCCTTTCTCC
2161 CTTGCGGAAG CGTGGCGCTT TCTCAATGCT CACGCTGTAG GTATCTCAGT TCGGTGTAGG
2221 TCGTTCGCTC CAAGCTGGGC TGTGTGCACG AACCCCCCGT TCAGCCCGAC CGCTGCGCCT
2281 TATCCGGTAA CTATCGTCTT GAGTCCAACC CGGTAAGACA CGACTTATCG CCACTGGCAG
2341 CAGCCACTGG TAACAGGATT AGCAGAGCGA GGTATGTAGG CGGTGCTACA GAGTTCTTGA
2401 AGTGGTGGCC TAACTACGGC TACACTAGAA GGACAGTATT TGGTATCTGC GCTCTGCTGA
2461 AGCCAGTTAC CTTCCGAAAA AGAGTTGGTA GCTCTTGATC CGGCAAAACAA ACCACCGCTG
2521 GTAGCGGTGG TTTTTTTGTT TGCAAGCAGC AGATTACGCG CAGAAAAAAA GGATCTCAAG
2581 AAGATCCTTT GATCTTTTCT ACGGGGTCTG ACGCTCAGTG GAACGAAAC TCACGTTAAG
2641 GGATTTTGGT CATGCCATAA CTTGCTATAG CATACATTAT ACGAAGTTAT GGCATGAGAT
2701 TATCAAAAAG GATCTTCACC TAGATCCTTT TAAATTAAAA ATGAAGTTT AAATCAATCT
2761 AAAGTATATA TGAGTAACT TGGTCTGACA GTTACCAATG CTTAATCAGT GAGGCACCTA
2821 TCTCAGCGAT CTGTCTATTT CTTTCATCCA TAGTTGCTG ACTCCCCGTC GTGTAGATAA
2881 CTACGATACG GGAGGGCTTA CCATCTGGCC CCAAGTGCTGC AATGATACCG CGAGACCCAC
2941 GCTCACCAGC TCCAGATTTA TCAGCAATAA ACCAGCCAGC CGGAAGGGCC GAGCGCAGAA
3001 GTGGTCTCTG AACTTTATCC GCCTCCATCC AGTCTATTAA TTGTTGCCGG GAAGCTAGAG
3061 TAAGTAGTTC GCCAGTTAAT AGTTTGCGCA ACGTTGTTGC CATTGCTACA GGCATCGTGG
3121 TGTCACGCTC GTCGTTTGGT ATGGCTTCAT TCAGCTCCGG TTCCCAACGA TCAAGGCGAG-

Figure 48C

3181 TTACATGATC CCCCATGTTG TGCAAAAAAG CGGTTAGCTC CTTCGGTCCT CCGATCGTTG
3241 TCAGAAGTAA GTTGGCCGCA GTGTTATCAC TCATGGTTAT GGCAGCACTG CATAATTCTC
3301 TTAATGTCAT GCCATCCGTA AGATGCTTTT CTGTGACTGG TGAGTACTCA ACCAAGTCAT
3361 TCTGAGAATA GTGTATGCGG CGACCCAGTT GCTCTTGCCC GGCCTCAATA CGGGATAATA
3421 CCGCGCCACA TAGCAGAACT TTAAGAGTGC TCATCATTGG AAAACGTTCT TCGGGGCGAA
3481 AACTCTCAAG GATCTTACCG CTGTTGAGAT CCAGTTCGAT GTAACCCACT CGTGCAACCA
3541 ACTGATCTTC AGCATCTTTT ACTTTCACCA GCGTTTCTGG GTGAGCAAAA ACAGGAAGGC
3601 AAAATGCCGC AAAAAAGGGA ATAAGGGCGA CACGGAATG TTGAATACTC ATACTCTTCC
3661 TTTTCAATA TTATTGAAGC ATTTATCAGG GTTATTGTCT CATGCCAGGG GTGGGCACAC
3721 ATATTTGATA CCAGCGATCC CTACACAGCA CATAATTCAA TGCGACTTCC CTCTATCGCA
3781 CATCTTAGAC CTTTATTCTC CCTCCAGCAC ACATCGAAGC TGCCGAGCAA GCCGTTCTCA
3841 CCAGTCCAAG ACCTGGCATG AGCGGATACA TATTTGAATG TATTTAGAAA AATAACAAA
3901 TAGGGGTTCC GCGCACATT CCCCAGAAAG TGCCACCTGA AATTGTAAAC GTTAATATTT
3961 TGTAAAAATT CGCGTTAAAT TTTTGTAAA TCAGCTCATT TTTTAACCA TAGGCCGAAA
4021 TCGGCAAAAT CCCTTATAAA TCAAAGAAT AGACCGAGAT AGGGTTGAGT GTTGTTCAG
4081 TTTGGAACAA GAGTCCACTA TTAAGAAGC TGGACTCCAA CGTCAAAGGG CGAAAAACCG
4141 TCTATCAGGG CGATGGCCCA CTACGTGAAC CATCACCTA ATCAAGTTTT TTGGGGTCGA
4201 GGTGCCGTAA AGCACTAAAT CGGAACCCTA AAGGGAGCCC CCGATTTAGA GCTTGACGGG
4261 GAAAGCCGGC GAACGTGGCG AGAAGGAAG GGAAGAAAGC GAAAGGAGCG GCGCTAGGG
4321 CGCTGGCAAG TGTAGCGGC ACGCTGCGCG TAACCACCAC ACCCGCCGCG CTTAATGCGC
4381 CGCTACAGGG CGCGTC

FIGURE 48D

144/240



145/240

pDONR201 4470 bp (rotated to position 3516)

<u>Location (Base Nos.)</u>		<u>Gene Encoded</u>
260..29		attP1
656..961		ccdB
1099..1184		ccdA
1303..1962		CmR
2210..2442		attP2
2565..3374		Kmr
3495..4134		ori

1	GTAAACGCTA	GCATGGATCT	CGGGCCCCAA	ATAATGATTT	TATTTTGACT	GATAGTGACC
61	TGTTCTGTTG	AACAAATTGA	TGAGCAATGC	TTTTTTATAA	TGCCAACTTT	GTACAAAAAA
121	GCTGAACGAG	AAACGTAAAA	TGATATAAAT	ATCAATATAT	TAAATTAGAT	TTGTCATAAA
181	AAACAGACTA	CATAATACTG	TAAAACACAA	CATATCCAGT	CACTATGAAT	CAACTACTTA
241	GATGGTATTA	GTGACCTGTA	GTGACCGGAC	AGCCTTCCAA	ATGTTCTTCG	GGTGATGCTG
301	CCAACCTAGT	CGACCGACAG	CCTTCCAAAT	GTTCTTCTCA	AACGGAATCG	TCGTATCCAG
361	CCTACTCGCT	ATTGTCCTCA	ATGCCGTATT	AAATCATAAA	AAGAAATAAG	AAAAAGAGGT
421	GCGAGCCTCT	TTTTTGTGTG	ACAAAATAAA	AACATCTACC	TATTCATATA	CGCTAGTGTC
481	ATAGTCCTGA	AAATCATCTG	CATCAAGAAC	AATTTCCAAA	CTCTTATACT	TTTCTCTTAC
541	AAGTCGTTCT	GCTTCATCTG	GATTTTCAGC	CTCTATACTT	ACTAAACGTG	ATAAAGTTTC
601	TGTAATTTCT	ACTGTATCGA	CCTGCAGACT	GGCTGTGTAT	AAGGGAGCCT	GACATTTATA
661	TTCCCCAGAA	CATCAGGTTA	ATGGCGTTTT	TGATGTCAAT	TTGCGGGTGG	CTGAGATCAG
721	CCACTTCTTC	CCCGATAACG	GAGACCGGCA	CACTGGCCAT	ATCGGTGGTC	ATCATGCGCC
781	AGCTTTTCATC	CCCGATATGC	ACCACCGGGT	AAAGTTCACG	GGAGACTTTA	TCTGACAGCA
841	GACGTGCACT	GGCCAGGGGG	ATCACCATCC	GTGCCCCGGG	CGTGTCAATA	ATATCACTCT
901	GTACATCCAC	AAACAGACGA	TAACGGCTCT	CTCTTTTATA	GGTGTAACC	TTAAACTGCA
961	TTTACCAGT	CCCTGTTCTC	GTCAGCAAAA	GAGCCGTTCA	TTTCAATAAA	CCGGGCGACC
1021	TCAGCCATCC	CTTCTGATT	TTCCGCTTTC	CAGCGTTCGG	CACGCAGACG	ACGGGCTTCA
1081	TTCTGCTATG	TTGTGCTTAC	CAGACCGGAG	ATATTGACAT	CATATATGCC	TTGAGCAACT
1141	GATAGCTGTC	GCTGTCAACT	GTCACTGTAA	TACGCTGCTT	CATAGCACAC	CTCTTTTGA
1201	CATACTTCGG	GTATACATAT	CAGTATATAT	TCTTATACCG	CAAAAATCAG	CGCGCAATA
1261	CGCATACTGT	TATCTGGCTT	TTAGTAAGCC	GGATCCACGC	GATTACGCCC	CGCCCTGCCA
1321	CTCATCGCAG	TACTGTGTGA	ATTCAATTAAG	CATTCTGCCG	ACATGGAAGC	CATCACAGAC
1381	GGCATGATGA	ACCTGAATCG	CCAGCGGCAT	CAGCACCTTG	TCGCCTTGCG	TATAATATTT
1441	GCCCATGGTG	AAAACGGGGG	CGAAGAAGTT	GTCCATATTG	GCCACGTTTA	AATCAAAACT
1501	GGTGAAGTC	ACCCAGGGAT	TGGCTGAGAC	GAAAAACATA	TTCTCAATAA	ACCCTTTAGG
1561	GAAATAGGCC	AGGTTTTCAC	CGTAACACGC	CACATCTTGC	GAATATATGT	TAGAGAACTG
1621	CCGGAATCG	TCGTGGTATT	CACTCCAGAG	CGATGAAAAC	GTTTCAGTTT	GCTCATGGAA
1681	AACGGTGTAA	CAAGGGTGAA	CACTATCCCA	TATCACCAGC	TCACCGTCTT	TCATTGCCAT
1741	ACGGAATTCC	GGATGAGCAT	TCATCAGGCG	GGCAAGAATG	TGAATAAAGG	CCGGATAAAA
1801	CTTGTGCTTA	TTTTTCTTTA	CGGTCTTTAA	AAAGGCCGTA	ATATCCAGCT	GAACGGTCTG
1861	GTTATAGGTA	CATTGAGCAA	CTGACTGAAA	TGCCTCAAAA	TGTTCTTTAC	GATGCCATTG
1921	GGATATATCA	ACGGTGGTAT	ATCCAGTGAT	TTTTTTCTCC	ATTTTAGCTT	CCTTAGCTCC
1981	TGAAATCTC	GATAACTCAA	AAAATACGCC	CGGTAGTGAT	CTTATTTTCA	TATGGTGAAA
2041	GTTGGAACCT	CTTACGTGCC	GATCAACGTC	TCATTTTCGC	CAAAAGTTGG	CCCAGGGCTT
2101	CCCGGTATCA	ACAGGGACAC	CAGGATTTAT	TTATTCTGCG	AAGTGATCTT	CCGTCACAGG
2161	TATTTATTTC	GCGCAAAGTG	CGTCGGGTGA	TGCTGCCAAC	TTAGTCGACT	ACAGGTCACT
2221	AATACCATCT	AAGTAGTTGA	TTCATAGTGA	CTGGATATGT	TGTGTTTTAC	AGTATTATGT
2281	AGTCTGTTTT	TTATGCAAAA	TCTAATTTAA	TATATTGATA	TTTATATCAT	TTTACGTTTC
2341	TCGTTTCAGCT	TTCTTGTA	AAGTTGGCAT	TATAAGAAAG	CATTGCTTAT	CAATTTGTTG
2401	CAACGAACAG	GTCACATCA	GTCAAAATAA	AATCATTATT	TGCCATCCAG	CTGCAGCTCT
2461	GGCCCGTGTC	TCAAAATCTC	TGATGTTACA	TTGCACAAGA	TAAAAATATA	TCATCATGAA
2521	CAATAAAACT	GTCTGCTTAC	ATAAACAGTA	ATACAAGGGG	TGTTATGAGC	CATATTCAAC
2581	GGGAAACGTC	GAGGCCGCGA	TTAAATTCCA	ACATGGATGC	TGATTTATAT	GGGTATAAAT
2641	GGGCTCGCGA	TAATGTCGGG	CAATCAGGTG	CGACAATCTA	TCGCTTGAT	GGGAAGCCCG
2701	ATGCGCCAGA	GTTGTTTCTG	AAACATGGCA	AAGGTAGCGT	TGCCAATGAT	GTTACAGATG -

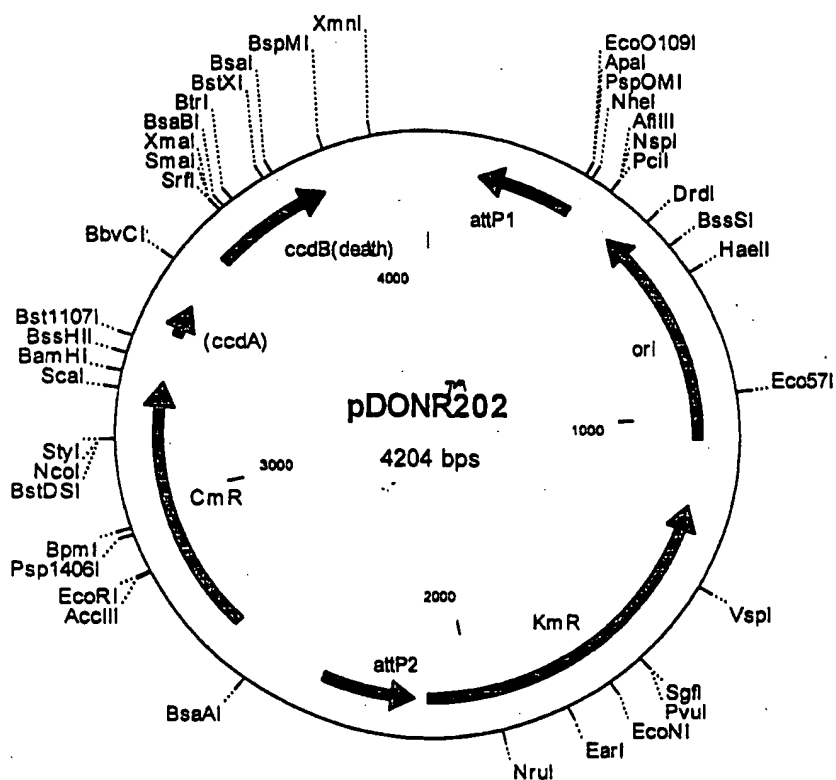
FIGURE 49B

146/240

2761 AGATGGTCAG ACTAAACTGG CTGACGGAAT TTATGCCTCT TCCGACCATC AAGCATTTTA
2821 TCCGTACTCC TGATGATGCA TGGTTACTCA CCACTGCGAT CCCC GGAAAA ACAGCATTC
2881 AGGTATTAGA AGAATATCCT GATTTCAGGTG AAAATATTGT TGATGCGCTG GCAGTGTTCC
2941 TGC GCGGTT GCATTCGATT CCTGTTTGTA ATTGTCCTTT TAACAGCGAT CGCGTATTTT
3001 GTCTCGCTCA GCGCGAATCA CGAATGAATA ACGGTTTGGT TGATGCGAGT GATTTTGATG
3061 ACGAGCGTAA TGGCTGGCCT GTTGAACAAG TCTGGAAAGA AATGCATAAA CTTTGGCCAT
3121 TCTCACC GGA TTCAGTCGTC ACTCATG GTG ATTTCTCACT TGATAACCTT ATTTTGTACG
3181 AGGGGAAATT AATAGGTTGT ATTGATGTTG GACGAGTCGG AATCGCAGAC CGATACCAGG
3241 ATCTTGCCAT CCTATGGAAC TGCCTCGGTG AGTTTTCTCC TTCATTACAG AAACGGCTTT
3301 TTCAAAAATA TGGTATTGAT AATCCTGATA TGAATAAATT GCAGTTTCAT TTGATGCTCG
3361 ATGAGTTTTT CTAATCAGAA TTGGTTAATT GGTGTAACA CTGGCAGAGC ATTACGCTGA
3421 CTTGACGGGA CGGCGCAAGC TCATGACCAA AATCCCTTAA CGTGAGTTTT CGTTCCTACTG
3481 AGCGTCAGAC CCCGTAGAAA AGATCAAAGG ATCTTCTTGA GATCCTTTTT TTCTGCGCGT
3541 AATCTGCTGC TTGCAACAA AAAAACCCACC GCTACCAGCG GTGGTTTGTT TGCCGGATCA
3601 AGAGCTACCA ACTCTTTTTT CGAAGGTAAC TGGCTTCAGC AGAGCGCAGA TACCAAATAC
3661 TGTCCTTCTA GTGTAGCCGT AGTTAGGCCA CCACTTCAAG AACTCTGTAG CACCGCCTAC
3721 ATACCTCGCT CTGCTAATCC TGTTACCAGT GGCTGCTGCC AGTGCGGATA AGTCGTGTCT
3781 TACCGGGTTG GACTCAAGAC GATAGTTACC GGATAAGGCG CAGCGGTCGG GCTGAACGGG
3841 GGGTTCGTGC ACACAGCCCA GCTTGGAGCG AACGACCTAC ACCGAACTGA GATACCTACA
3901 GCGTGAGCTA TGAGAAAGCG CCACGCTTCC CGAAGGGAGA AAGGCGGACA GGTATCCGGT
3961 AAGCGGCAGG GTCGGAACAG GAGAGCGCAC GAGGGAGCTT CCAGGGGGAA ACGCCTGGTA
4021 TCTTTATAGT CCTGTCGGGT TTCGCCACCT CTGACTTGAG CGTCGATTTT TGTGATGCTC
4081 GTCAGGGGGG CGGAGCCTAT GGAAAAACGC CAGCAACGCG GCCTTTTTAC GGTTCCTGGC
4141 CTTTTGCTGG CTTTTTGCTC ACATGTTCTT TCCTGCGTTA TCCCCTGATT CTGTGGATAA
4201 CCGTATTACC GCTAGCCAGG AAGAGTTTGT AGAAACGCAA AAAGGCCATC CGTCAGGATG
4261 GCCTTCTGCT TAGTTTGATG CCTGGCAGTT TATGGCGGGC GTCTGCCCCG CCACCCTCCG
4321 GGCCGTTGCT TCACAACGTT CAAATCCGCT CCCGGCGGAT TTGTCCTACT CAGGAGAGCG
4381 TTCACCGACA AACCAACAGAT AAAACGAAAG GCCCAGTCTT CCGACTGAGC CTTTCGTTTT
4441 ATTTGATGCC TGGCAGTTCC CTA CTCTCTCGC

FIGURE 49C

147/240
FIGURE 50A: pDONR202 (kan^R)



148/240

pDONR202 4204 bp

<u>Location (Base Nos.)</u>			<u>Gene Encoded</u>			
369..127			attP1			
486..1059			ori			
1228..2107			KmR			
2381..2140			attP2			
2629..3288			CmR			
3408..3492			inactivated ccdA			
3630..3935			ccdB			
1	CGGCATTGAG	GACAATAGCG	AGTAGGCTGG	ATACGACGAT	TCCGTTTGAG	AAGAACATTT
61	GGAAGGCTGT	CGGTCGACTA	AGTTGGCAGC	ATCACC CGAA	GAACATTTGG	AAGGCTGTGC
121	GTCGACTACA	GGTCACTAAT	ACCATCTAAG	TAGTTGATTC	ATAGTGACTG	GATATGTTGT
181	GTTTTACAGT	ATTATGTAGT	CTGTTTTTTA	TGCAAAATCT	AATTTAATAT	ATTGATATTT
241	ATATCATTTT	ACGTTTCTCG	TTCAGCTTTT	TTGTACAAAG	TTGGCATTAT	AAAAAAGCAT
301	TGCTCATCAA	TTTGTGCAA	CGAACAGGTC	ACTATCAGTC	AAAATAAAAT	CATTATTTGG
361	GGCCCCGAGAT	CCATGCTAGC	GGTAATACGG	TTATCCACAG	AATCAGGGGA	TAACGCAGGA
421	AAGAACATGT	GAGCAAAAGG	CCAGCAAAAG	GCCAGGAACC	GTAAAAGGC	CGCGTTGCTG
481	GCGTTTTTCC	ATAGGCTCCG	CCCCCTGAC	GAGCATCACA	AAAATCGACG	CTCAAGTCAG
541	AGGTGGCGAA	ACCCGACAGG	ACTATAAAGA	TACCAGGCGT	TTCCCCCTGG	AAGCTCCCTC
601	GTGCGCTCTC	CTGTTCCGAC	CCTGCCGCTT	ACCGGATACC	TGTCCGCCTT	TCTCCCTTCG
661	GGAAGCGTGG	CGCTTTCTCA	TAGCTCACGC	TGTAGGTATC	TCAGTTCGGT	GTAGGTCGTT
721	CGCTCCAAGC	TGGGCTGTGT	GCACGAACCC	CCCGTTCAGC	CCGACCGCTG	CGCCTTATCC
781	GGTAACTATC	GTCTTGAGTC	CAACCCGGTA	AGACACGACT	TATCGCCACT	GGCAGCAGCC
841	ACTGGTAACA	GGATTAGCAG	AGCGAGGTAT	GTAGGCGGTG	CTACAGAGTT	CTTGAAGTGG
901	TGGCCTAACT	ACGGCTACAC	TAGAAGGACA	GTATTTGGTA	TCTGCGCTCT	GCTGAAGCCA
961	GTTACCTTCG	GAAAAAGAGT	TGGTAGCTCT	TGATCCGGCA	AACAAACCAC	CGCTGGTAGC
1021	GGTGGTTTTT	TTGTTTGCAA	GCAGCAGATT	ACGCGCAGAA	AAAAAGGATC	TCAAGAAGAT
1081	CCTTTGATCT	TTTCTACGGG	GTCTGACGCT	CAGTGGAAACG	AAAACCTCAG	TTAAGGGATT
1141	TTGGTCATGA	GCTTGCGCCG	TCCCGTCAAG	TCAGCGTAAT	GCTCTGCCAG	TGTTACAACC
1201	AATTAACCAA	TTCTGATTAG	AAAAACTCAT	CGAGCATCAA	ATGAAACTGC	AATTTATTCA
1261	TATCAGGATT	ATCAATACCA	TATTTTTGAA	AAAGCCGTTT	CTGTAATGAA	GGAGAAAACCT
1321	CACCGAGGCA	GTTCCATAGG	ATGGCAAGAT	CCTGGTATCG	GTCTGCGATT	CCGACTCGTC
1381	CAACATCAAT	ACAACCTATT	AATTTCCCTT	CGTCAAAAAT	AAGGTTATCA	AGTGAGAAAT
1441	CACCATGAGT	GACGACTGAA	TCCGGTGAGA	ATGGCAAAAG	TTTATGCATT	TCTTTCCAGA
1501	CTTGTTCAAC	AGGCCAGCCA	TTACGCTCGT	CATCAAAATC	ACTCGCATCA	ACCAAACCGT
1561	TATTCTATCG	TGATTGCGCC	TGAGCGAGAC	GAAATACGCG	ATCGCTGTTA	AAAGGACAAT
1621	TACAAACAGG	AATCGAATGC	AACCGGCGCA	GGAACACTGC	CAGCGCATCA	ACAATATTTT
1681	CACCTGAATC	AGGATATTCT	TCTAATACCT	GGAATGCTGT	TTTCCGGGG	ATCGCAGTGG
1741	TGAGTAACCA	TGCATCATCA	GGAGTACGGA	TAAAATGCTT	GATGGTCGGA	AGAGGCATAA
1801	ATTCGCTCAG	CCAGTTTAGT	CTGACCATCT	CATCTGTAAC	ATCATTTGGA	ACGCTACCTT
1861	TGCCATGTTT	CAGAAACAAC	TCTGGCGCAT	CGGGCTTCCC	ATACAAGCGA	TAGATTGTCG
1921	CACCTGATTG	CCCGACATTA	TCGCGAGCCC	ATTTATACCC	ATATAAATCA	GCATCCATGT
1981	TGGAATTTAA	TCGCGGCCTC	GACGTTTCCC	GTTGAATATG	GCTCATAACA	CCCCTTGAT
2041	TACTGTTTAT	GTAAGCAGAC	AGTTTTATTG	TTCATGATGA	TATATTTTTA	TCTTGTCGAA
2101	TGTAACATCA	GAGATTTTGA	GACACGGGCC	AGAGCTGCAG	CTGGATGGCA	AATAATGATT
2161	TTATTTTGAC	TGATAGTGAC	CTGTTTCGTT	CAACAAATTG	ATAAGCAATG	CTTCTTATA
2221	ATGCCAACTT	TGTACAAGAA	AGCTGAACGA	GAAACGTAAA	ATGATATAAA	TATCAATATA
2281	TTAAATTAGA	TTTTGCATAA	AAAACAGACT	ACATAATACT	GTAAAACACA	ACATATCCAG
2341	TCACATGAA	TCACTACTT	AGATGGTATT	AGTGACCTGT	AGTCGACTAA	GTTGGCAGCA
2401	TCACCCGACG	CACTTTGCGC	CGAATAAATA	CCTGTGACGG	AAGATCACTT	CGCAGAATAA
2461	ATAAATCCTG	GTGTCCCTGT	TGATACCGGG	AAGCCCTGGG	CCAACCTTTG	CGGAAAATGA
2521	GACGTTGATC	GGCACGTAAG	AGGTTCCAAC	TTTCACCATA	ATGAAATAAG	ATCACTACCG
2581	GGCGTATTTT	TTGAGTTATC	GAGATTTTCA	GGAGCTAAGG	AAGCTAAAAT	GGAGAAAAAA
2641	ATCACTGGAT	ATACCACCGT	TGATATATCC	CAATGGCATC	GTAAAGAACA	TTTTGAGGCA
2701	TTTCAGTCAG	TTGCTCAATG	TACCTATAAC	CAGACCGTTC	AGCTGGATAT	TACGGCCTTT

FIGURE 50B

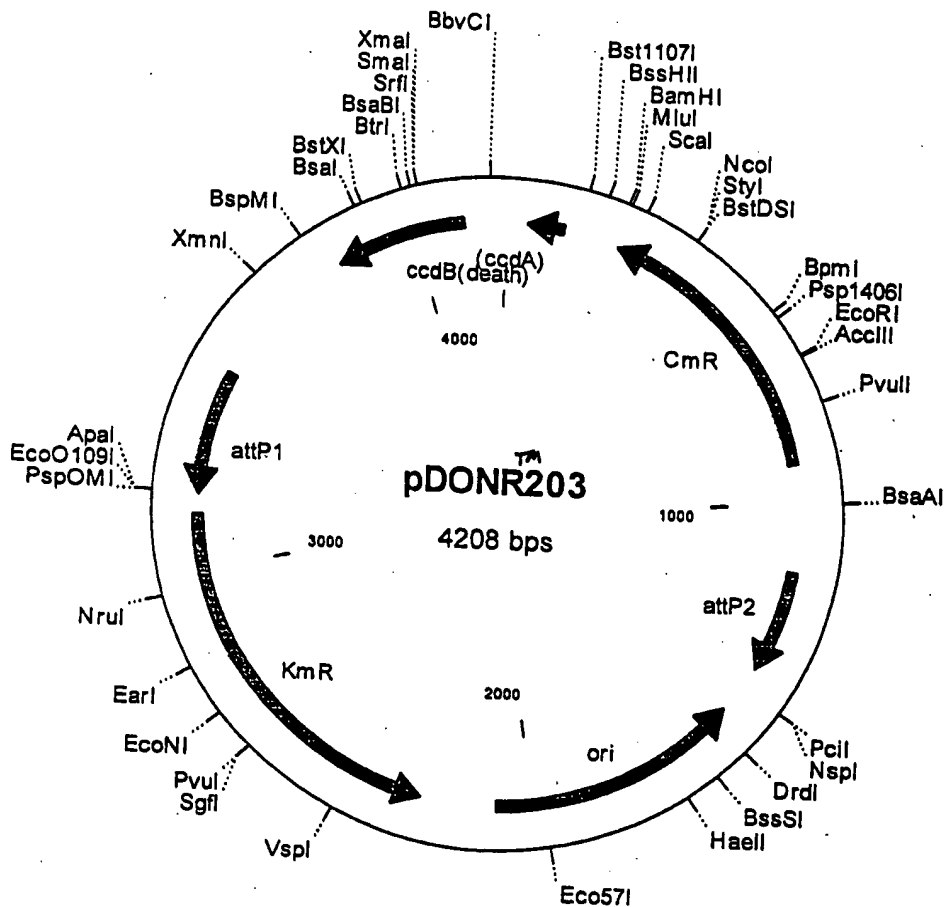
149/240

2761 TTAAAGACCG TAAAGAAAA TAAGCACAAG TTTTATCCGG CCTTTATTCA CATTCTTGCC
2821 CGCCTGATGA ATGCTCATCC GGAATTCGT ATGGCAATGA AAGACGGTGA GCTGGTGATA
2881 TGGGATAGTG TTCACCCTTG TTACACCGTT TTCCATGAGC AAAGTGAAC GTTTTCATCG
2941 CTCTGGAGTG AATACCACGA CGATTTCGG CAGTTTCTAC ACATATATTC GCAAGATGTG
3001 GCGTGTTACG GTGAAAACCT GGCCTATTTT CCTAAAGGGT TTATTGAGAA TATGTTTTTC
3061 GTCTCAGCCA ATCCCTGGGT GAGTTTCACC AGTTTGTATT TAAACGTGGC CAATATGGAC
3121 AACTTCTTCG CCCCCGTTTT CACCATGGGC AAATATTATA CGCAAGGCGA CAAGGTGCTG
3181 ATGCCGCTGG CGATTACGGT TCATCATGCC GTCTGTGATG GCTTCCATGT CGGCAGAATG
3241 CTTAATGAAT TACAACAGTA CTGCGATGAG TGGCAGGGCG GGGCGTAATC GCGTGGATCC
3301 GGCTTACTAA AAGCCAGATA ACAGTATGCG TATTTGCGCG CTGATTTTTG CGGTATAAGA
3361 ATATATACTG ATATGTATAC CCGAAGTATG TCAAAAAGAG GTGTGCTATG AAGCAGCGTA
3421 TTACAGTGAC AGTTGACAGC GACAGCTATC AGTTGCTCAA GGCATATATG ATGTCAATAT
3481 CTCCGGTCTG GTAAGCACAA CCATGCAGAA TGAAGCCCGT CGTCTGCGTG CCGAACGCTG
3541 GAAAGCGGAA AATCAGGAAG GGATGGCTGA GGTGCGCCGG TTTATTGAAA TGAACGGCTC
3601 TTTTGCTGAC GAGAACAGGG ACTGGTGAAA TGCAGTTTAA GGTTTACACC TATAAAGAG
3661 AGAGCCGTTA TCGTCTGTTT GTGGATGTAC AGAGTGATAT TATTGACACG CCCGGGCGAC
3721 GGATGGTGAT CCCCCTGGCC AGTGCACGTC TGCTGTCAGA TAAAGTCTCC CGTGAACTTT
3781 ACCCGGTGGT GCATATCGGG GATGAAAGCT GGCGCATGAT GACCACCGAT ATGGCCAGTG
3841 TGCCGGTCTC CGTTATCGGG GAAGAAGTGG CTGATCTCAG CCACCGCGAA AATGACATCA
3901 AAAACGCCAT TAACCTGATG TTCTGGGGAA TATAAATGTC AGGCTCCCTT ATACACAGCC
3961 AGTCTGCAGG TCGATACAGT AGAAATTACA GAACTTTAT CACGTTTAGT AAGTATAGAG
4021 GCTGAAAATC CAGATGAAGC CGAACGACTT GTAAGAGAAA AGTATAAGAG TTGTGAAATT
4081 GTTCTTGATG CAGATGATT TCAGGACTAT GACACTAGCG TATATGAATA GGTAGATGTT
4141 TTTATTTTGT CACACAAAA AGAGGCTCGC ACCTCTTTTT CTTATTTCTT TTTATGATTT
4201 AATA

FIGURE 50C

150/240

FIGURE 51A pDONR203 (kan^R)



pDONR203 4208 bp

<u>Location (Base Nos.)</u>	<u>Gene Encoded</u>
47..131	inactivated ccdA
251..910	CmR
1158..1398	attP2
1509..2082	ori
2251..3130	KmR
3464..3174	attP1
3812..4117	ccdB

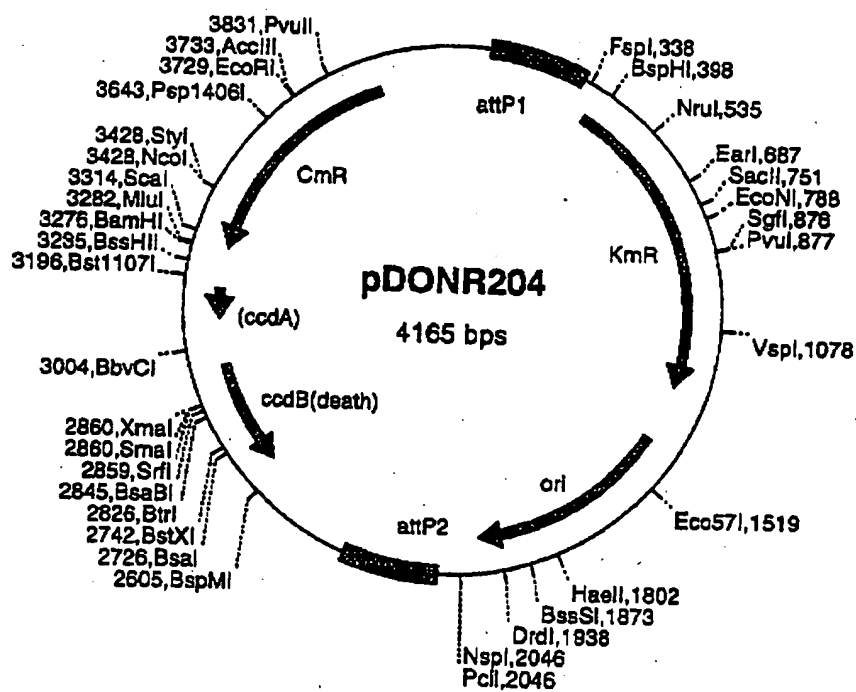
```

1 GCGTTCGGCA CGCAGACGAC GGGCTTCATT CTGCATGGTT GTGCTTACCA GACCGGAGAT
61 ATTGACATCA TATATGCCTT GAGCAACTGA TAGCTGTCGC TGTCAACTGT CACTGTAATA
121 CGCTGCTTCA TAGCACACCT CTTTTTGACA TACTTCGGGT ATACATATCA GTATATATTC
181 TTATACCGCA AAAATCAGCG CGCAAATACG CATACTGTTA TCTGGCTTTT AGTAAGCCGG
241 ATCCACGCGT TTACGCCCCG CCTGCCACT CATCGCAGTA CTGTTGTAAT TCATTAAGCA
301 TTCTGCCGAC ATGGAAGCCA TCACAGACGG CATGATGAAC CTGAATCGCC AGCGGCATCA
361 GCACCTTGTC GCCTTGCGTA TAATATTTGC CCATGGTGAA AACGGGGGCG AAGAAGTTGT
421 CCATATTGGC CACGTTTAAA TCAAACTGG TGAACCTCAC CCAGGGATTG GCTGAGACGA
481 AAAACATATT CTCAATAAAC CCTTTAGGGA AATAGGCCAG GTTTTCAACCG TAACACGCCA
541 CATCTTGCGA ATATATGTGT AGAACTGCC GGAAATCGTC GTGGTATTCA CTCCAGAGCG
601 ATGAAAACGT TTCAGTTTGC TCATGGAAAA CGGTGTAACA AGGGTGAACA CTATCCCAT
661 TCACCAGCTC ACCGTCTTTC ATTGCCATAC GGAATTCGG ATGAGCATTC ATCAGGCGGG
721 CAAGAATGTG AATAAAGGCC GGATAAAACT TGTGCTTATT TTTCTTTACG GTCTTTAAAA
781 AGGCCGTAAT ATCCAGCTGA ACGGTCTGGT TATAGGTACA TTGAGCAACT GACTGAAATG
841 CCTCAAAATG TTCTTTACGA TGCCATTGGG ATATATCAAC GGTGGTATAT CCAGTGATTT
901 TTTTCTCCAT TTTAGCTTCC TTAGCTCCTG AAAATCTCGA TAACTCAAAA AATACGCCCG
961 GTAGTGATCT TATTTCAATTA TGGTGAAAGT TGGAACCTCT TACGTGCCGA TCAACGTCTC
1021 ATTTTCGCCA AAAGTTGGCC CAGGGCTTCC CGGTATCAAC AGGGACACCA GGATTTATTT
1081 ATTCTGCGAA GTGATCTTCC GTCACAGGTA TTTATTCGGC GCAAAGTGCG TCGGGTGATG
1141 CTGCCAACTT AGTCGACTAC AGGTCACTAA TACCATCTAA GTAGTTGATT CATAGTGACT
1201 CGATATGTTG TGTTTTACAG TATTATGTAG TCTGTTTTTT ATGCAAAATC TAATTTAATA
1261 TATTGATATT TATATCATTT TACGTTTCTC GTTCAGCTTT CTTGTACAAA GTTGGCATT
1321 TAAGAAAGCA TTGCTTATCA ATTTGTTGCA ACGAACAGGT CACTATCAGT CAAAATAAAA
1381 TCATTATTTG CCATCCAGCT AGCGGTAATA CGGTATCCA CAGAATCAGG GGATAACGCA
1441 GGAAAGAACA TGTGAGCAAA AGGCCAGCAA AAGGCCAGGA ACCGTAAAAA GGCCGCGTTG
1501 CTGGCGTTTT TCCATAGGCT CCGCCCCCCT GACGAGCATC ACAAATTCG ACGCTCAAGT
1561 CAGAGGTGGC GAAACCCGAC AGGACTATAA AGATACCAGG CGTTTCCCCC TGGAAGCTCC
1621 CTCGTGCGCT CTCCTGTTC GACCCTGCCG CTTACCGGAT ACCTGTCCGC CTTTCTCCCT
1681 TCGGGAAGCG TGGCGCTTTC TCATAGCTCA CGCTGTAGGT ATCTCAGTTC GGTGTAGGTC
1741 GTTCGCTCCA AGCTGGGCTG TGTGCACGAA CCCCCGTTT AGCCCGACCG CTGCGCCTTA
1801 TCCGGTAACT ATCGTCTTGA GTCCAACCCG GTAAGACACG ACTTATCGCC ACTGGCAGCA
1861 GCCACTGGTA ACAGGATTAG CAGAGCGAGG TATGTAGGCG GTGCTACAGA GTTCTTGAAG
1921 TGTGCGCCTA ACTACGGCTA CACTAGAAGA ACAGTATTTG GTATCTGCGC TCTGCTGAAG
1981 CCAGTTACCT TCGGAAAAAG AGTTGGTAGC TCTTGATCCG GCAAACAAAC CACCGCTGGT
2041 AGCGGTGGTT TTTTGTGTTG CAAGCAGCAG ATTACGCGCA GAAAAAAGG ATCTCAAGAA
2101 GATCCTTTGA TCTTTTCTAC GGGGTCTGAC GCTCAGTGGA ACGAAAACTC ACGTTAAGGG
2161 ATTTTGGTCA TGAGCTTGCG CCGTCCCGTC AAGTCAGCGT AATGCTCTGC CAGTGTTACA
2221 ACCAATTAAC CAATTCTGAT TAGAAAACT CATCGAGCAT CAAATGAAAC TGCAATTTAT
2281 TCATATCAGG ATTATCAATA CCATATTTT GAAAAAGCCG TTTCTGTAAT GAAGGAGAAA
2341 ACTCACCGAG GCAGTTCCAT AGGATGSCAA GATCCTGGTA TCGGTCTGCG ATTCCGACTC
2401 GTCCAACATC AATACAACCT ATTAATTTCC CCTCGTCAAA AATAAGGTTA TCAAGTGAGA
2461 AATCACCATG AGTGACGACT GAATCCGGTG AGAATGGCAA AAGTTTATGC ATTTCTTTCC
2521 AGACTTGTTT AACAGGCCAG CCATTACGCT CGTCATCAAA ATCACTCGCA TCAACCAAAC
2581 CGTTATTATC TCGTGATTGC GCCTGAGCGA GACGAAATAC GCGATCGCTG TTAAGAGGAC
2641 AATTACAAAC AGGAATCGAA TGCAACCGGC GCAGGAACAC TGCCAGCGCA TCAACAATAT
2701 TTTACCTGA ATCAGGATAT TCTTCTAATA CCTGGAATGC TGTTTTTCG GGGATCGCAG-

```

2761 TGGTGAGTAA CCATGCATCA TCAGGAGTAC GGATAAAATG CTTGATGGTC GGAAGAGGCA
2821 TAAATTCCGT CAGCCAGTTT AGTCTGACCA TCTCATCTGT AACATCATTG GCAACGCTAC
2881 CTTTGCCATG TTTCAGAAAC AACTCTGGCG CATCGGGCTT CCCATACAAG CGATAGATTG
2941 TCGCACCTGA TTGCCCCGACA TTATCGCGAG CCCATTTATA CCCATATAAA TCAGCATCCA
3001 TGTTGGAATT TAATCGCGGC CTCGACGTTT CCCGTTGAAT ATGGCTCATA ACACCCCTTG
3061 TATTACTGTT TATGTAAGCA GACAGTTTTA TTGTTTCATGA TGATATATTT TTATCTTGTT
3121 CAATGTAACA TCAGAGATTT TGAGACACGG GCCAGAGCTG CAGCTAGCAT GGATCTCGGG
3181 CCCCAAATAA TGATTTTATT TTGACTGATA GTGACCTGTT CGTTGCAACA AATTGATGAG
3241 CAATGCTTTT TTATAATGCC AACTTTGTAC AAAAAAGCTG AACGAGAAAC GTAAATGAT
3301 ATAAATATCA ATATATTAAT TTAGATTTTG CATAAAAAAC AGACTACATA ATACTGTAAA
3361 ACACAACATA TCCAGTCACT ATGAATCAAC TACTTAGATG GTATTAGTGA CCTGTAGTCG
3421 ACCGACAGCC TTCAAATGT TCTTCGGGTG ATGCTGCCAA CTTAGTCGAC CGACAGCCTT
3481 CCAAATGTTT TTCTCAAACG GAATCGTCGT ATCCAGCCTA CTCGCTATTG TCCTCAATGC
3541 CGTATTAAAT CATAAAAAGA AATAAGAAAA AGAGGTGCGA GCCTCTTTTT TGTGTGACAA
3601 AATAAAAACA TCTACCTATT CATATACGCT AGTGTCATAG TCCTGAAAAT CATCTGCATC
3661 AAGAACAATT TCACAACCTT TATACTTTTC TCTTACAAGT CGTTCGGCTT CATCTGGATT
3721 TTCAGCCTCT ATACTTACTA AACGTGATAA AGTTTCTGTA ATTTCTACTG TATCGACCTG
3781 CAGACTGGCT GTGTATAAGG GAGCCTGACA TTTATATTCC CCAGAACATC AGGTTAATGG
3841 CGTTTTTGAT GTCATTTTCG CGGTGGCTGA GATCAGCCAC TTCTTCCCCG ATAACGGAGA
3901 CCGGCACACT GGCCATATCG GTGGTCATCA TGGCCAGCT TTCATCCCCG ATATGCACCA
3961 CCGGGTAAAG TTCACGGGAG ACTTTATCTG ACAGCAGACG TGCACCTGGC AGGGGGATCA
4021 CCATCCGTCG CCCGGGCGTG TCAATAATAT CACTCTGTAC ATCCACAAAC AGACGATAAC
4081 GGCTCTCTCT TTTATAGGTG TAAACCTTAA ACTGCATTTT ACCAGTCCCT GTTCTCGTCA
4141 GCAAAAGAGC CGTTCATTTT AATAAACCGG GCGACCTCAG CCATCCCTTC CTGATTTTCC
4201 GCTTTCCA

FIGURE 51C

FIGURE 52A pDONR204 (kan^R)

154/240

pDONR204 4165 bp

```
1 CGGCATTGAG GACAATAGCG AGTAGGCTGG ATACGACGAT TCCGTTTGAG AAGAACATTT
61 GGAAGGCTGT CGGTCGACTA CAGGTCACCTA ATACCATCTA AGTAGTTGAA TCATAGTGAC
121 TGGATATGTT GTGTTTACCA GTATTATGTA GTCTGTTTTT TATGCAAAAT CTAATTTAAT
181 ATATTGATAT TTATATCATT TTACGTTTTCT CGTTCAGCTT TTTTGTACAA AGTTGGCATT
241 ATAAAAAGC ATTGCTTATC AATTTGTTGC AACGAACAGG TCACTATCAG TCAAAATAAA
301 ATCATTATTT GGGGCCCGAG ATCCATGCTA GCTGCAGTGC GCAGGGCCCC TGCTCTCAAAA
361 TCTCTGATGT TACATTGCAC AAGATAAAAA TATATCATCA TGAACAATAA AACTGTCTGC
421 TTACATAAAC AGTAATACAA GGGGTGTTAT GAGCCATATT CAACGGGAAA CGTCTTGCTG
481 GAGGCCGCGA TTAAATTCCA ACATGGATGC TGATTTATAT GGGTATAAAT GGGCTCGCGA
541 TAATGTCGGG CAATCAGGTG CGACAATCTT TCGATTGTAT GGGGAAGCCC ATGCGCCAGA
601 GTTGTTTCTG AAACATGGCA AAGGTAGCGT TGCCAATGAT GTTACAGATG AGATGGTCAG
661 ACTAAACCTG CTGACGGAAT TTATGCCTCT TCCGACCATC AAGCATTTTA TCCGTACTCC
721 TGATGATGCA TGGTTACTCA CCACTGCGAT CCGCGGGAAA ACAGCATTCC AGGTATTAGA
781 AGAATATCCT GATTCAAGTG AAAATATTGT TGATGCGCTG GCAGTGTTCG TCGCCGGGTT
841 GCATTGATTT CCTGTTTGTA ATTGTCCTTT TAACAGCGAT CGCGTATTTT GTCTCGCTCA
901 GCGCGAATCA CGAATGAATA ACGGTTTGGT TGATGCGAGT GATTTTGATG ACGAGCGTAA
961 TGGCTGGCCT GTTGAACAAG TCTGGAAAGA AATGCATACG CTTTTGCCAT TCTCACCGBA
1021 TTCAGTCGTC ACTCATGGTG ATTTCTCACT TGATAACCTT ATTTTGTACG AGGGGAAATT
1081 AATAGGTTGT ATTGATGTTG GACGAGTCGG AATCGCAGAC CGATACCAGG ATCTTGCCAT
1141 CCTATGGAAC TGCCTCGGTG AGTTTTCTCC TTCATTACAG AAACGGCTTT TTCAAAAATA
1201 TGGTATTGAT AATCCTGATA TGAATAAATT GCAGTTTCAT TTGATGCTCG ATGAGTTTTT
1261 CTAATCAGAA TTGGTTAATT GGTGTAACA CTGGCAGAGC ATTACGCTGA CTTGACGGGA
1321 CGGCGNCATG ACCAAAATCC CTTAACGTGA GTTTTCGTTT CACTGAGCGT CAGACCCCGT
1381 AGAAAAGATC AAAGGATCTT CTTGAGATCC TTTTTTCTG CGCGTAATCT GCTGCTTGCA
1441 AACAAAAAAA CCACCGCTAC CAGCGGTGGT TTGTTTGCCG GATCAAGAGC TACCAACTCT
1501 TTTTCCGAAG GTAACCTGGT TCAGCAGAGC GCAGATACCA AATACTGTCC TTCTAGTGTA
1561 GCGCTAGTTA GGCACCACT TCAAGAACTC TGTAGCACCG CCTACATACC TCGCTCTGCT
1621 AATCCTGTTA CCAGTGGCTG CTGCCAGTGG CGATAAGTCG TGTCTTACCG GGTGGACTC
1681 AAGACGATAG TTACCGGATA AGGCGCAGCG GTCGGGCTGA ACGGGGGGTT CGTGACACA
1741 GCCAGCTTG GAGCGAACGA CCTACACCGA ACTGAGATAC CTACAGCGTG AGCTATGAGA
1801 AAGCGCCACG CTTCCCGAAG GGAGAAAGGC GGACAGGTAT CCGGTAAGCG GCAGGGTCGG
1861 AACAGGAGAG CGCACGAGGG AGCTTCAGG GGGAAACGCC TGGTATCTTT ATAGTCTCTG
1921 CGGGTTTCGC CACCTCTGAC TTGAGCGCTG ATTTTGTGA TGCTCGTCAG GGGGGCGGAG
1981 CCTATGGAAG AACGCCAGCA ACGCGGCTT TTTACGGTTC CTGGCCTTTT GCTGGCCTTT
2041 TGCTCACATG TTCTTTCCTG CGTTATCCCC TGATTCTGTG GATAACCGTA TTACCGCTAG
2101 CTGGATCGGC AAATAATGAT TTTATTTTGA CTGATAGTGA CCTGTTCTGT GCAACAAATT
2161 GATAAGCAAT GCTTTTTTAT AATGCCAACT TTGTACAAGA AAGCTGAACG AGAAACGTAA
2221 AATGATATAA ATATCAATAT ATTAAATTAG ATTTTGCATA AAAAACAGAC TACATAATAC
2281 TGTAAAACAC AACATATCCA GTCACATGA TTCAACTACT TAGATGGTAT TAGTGACCTG
2341 TAGTCGACTA AGTTGGCAGC ATCACCAGAC GCACTTTGCG CCGAATAAAT ACCTGTGACG
2401 GAAGATCACT TCGCAGAATA AATAAATCCT GGTGTCCCTG TTGATACCGG GAAGCCCTGG
2461 GCCAACTTTT GCGGAAAATG AGACGTTGAT CGGCACATTT CACAACCTTT ATACTTTTCT
2521 CTTACAAGTC GTTCGGCTTC ATCTGGATTT TCAGCCTCTA TACTTACTAA ACGTGATAAA
2581 GTTCTGTAA TTTCTACTGT ATCGACCTGC AGACTGGCTG TGTATAACGG AGCCTGACAT
2641 TTATATTCCC CAGAACATCA GGTAAATGGC GTTTTGTATG TCATTTTCGC GGTGGCTGAG
2701 ATCAGCCACT TCTTCCCCGA TAACGGAGAC CGGCACACTG GCCATATCGG TGGTCATCAT
2761 GCGCCAGCTT TCATCCCCGA TATGCACCAC CGGGTAAAGT TCACGGGAGA CTTTATCTGA
2821 CAGCAGACGT GCACTGGCCA GGGGGATCAC CATCCGTCGC CCGGGCGTGT CAATAATATC
2881 ACTCTGTACA TCCACAAACA GACGATAACG GCTCTCTCTT TTATAGGTGT AAACCTTAAA
2941 CTGCATTTCA CCAGTCCCTG TTCTCGTCAG CAAAAGAGCC GTTCATTTCA ATAAACCGGG
3001 CGACCTCAGC CATCCCTTCC TGATTTTCCG CTTTCCAGCG TTCGGCACGC AGACGACGGG
3061 CTTCACTCTG CATGGTTGTG CTTACCAGAC CGGAGATATT GACATCATAT ATGCCTTGAG
3121 CAACTGATAG CTGTCGCTGT CAACTGTCAC TGTAATACGC TGCTTCATAG CACACCTCTT-
```

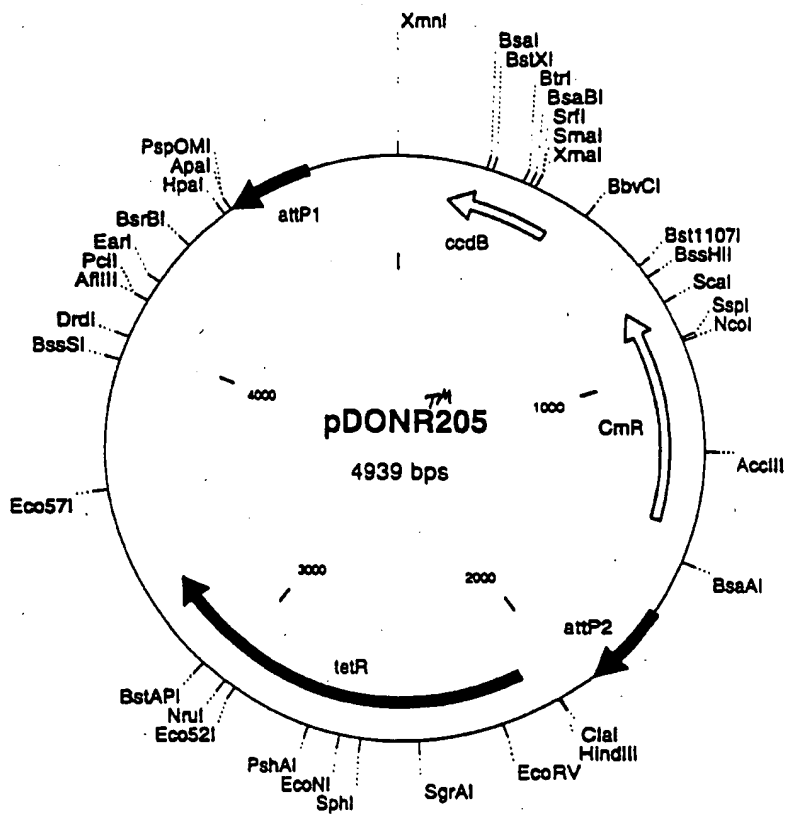
FIGURE 52B

3181 TTTGACATAC TTCGGGTATA CATATCAGTA TATATTCTTA TACCGCAAAA ATCAGCGCGC
3241 AAATACGCAT ACTGTTATCT GGCTTTTAGT AAGCCGGATC CACGCGTTTA CGCCCCGCCC
3301 TGCCACTCAT CGCAGTACTG TTGTAATTCA TTAAGCATTG TGCCGACATG GAAGCCATCA
3361 CAGACGGCAT GATGAACCTG AATCGCCAGC GGCATCAGCA CCTTGTCGCC TTGCGTATAA
3421 TATTTGCCCA TGGTGAAAAC GGGGGCGAAG AAGTTGTCCA TATTGGCCAC GTTTAAATCA
3481 AAACGGGTGA AACTCACCCA GGGATTGGCT GAGACGAAAA ACATATTCTC AATAAACCCCT
3541 TTAGGGAAAT AGGCCAGGTT TTCACCGTAA CACGCCACAT CTTGCGAATA TATGTGTAGA
3601 AACTGCCGGA AATCGTCGTG GTATTCACTC CAGAGCGATG AAAACGTTTC AGTTTGCTCA
3661 TGAAAACGG TGTAAACAAGG GTGAACACTA TCCCATATCA CCAGCTCACC GTCTTTCATT
3721 GCCATACGGA ATTCCGGATG AGCATTTCATC AGGCGGGCAA GAATGTGAAT AAAGGCCGGA
3781 TAAAACCTGT GCTTATTTTT CTTACGGTC TTTAAAAAGG CCGTAATATC CAGCTGAACG
3841 GTCTGGTTAT AGGTACATTG AGCAACTGAC TGAAATGCCT CAAAATGTTT TTTACGATGC
3901 CATTGGGATA TATCAACGGT GGTATATCCA GTGATTTTTT TCTCCATTTT AGCTTCCTTA
3961 GCTCCTGAAA ATCTCGATAA CTCAAAAAAT ACGCCCGGTA GTGATCTTAT TTCATTATGG
4021 TGAAAGTTGG AACCTCTTAC TGTTCTTGAT GCAGATGATT TTCAGGACTA TGACACTAGC
4081 ATATATGAAT AGGTAGATGT TTTTATTTTG TCACACAAAA AAGAGGCTCG CACCTCTTTT
4141 TCTTATTTCT TTTTATGATT TAATA

FIGURE 52C

156/240

Figure 53A: pDONR205 (tetR)



157/240

pDONR205 4939 bp

GGCATCAGCACCTTGTGCGCTTGCCTATAATATTTGCCCATGGTGAACCGGGGCGAAG
AAGTTGTCCATATTGGCCACGTTTAAATCAAACTGGTGAACCTCACCCAGGGATTGGCT
GAGACGAAAAACATATTCTCAATAAACCTTTAGGGAAATAGGCCAGGTTTTCACCGTAA
CACGCCACATCTTGCGAATATATGTGTAGAACTGCCGGAATCGTCGTGGTATTCACTC
CAGAGCGATGAAAACGTTTTCAGTTTGCTCATGGAACCGGTGTAACAAGGGTGAACACTA
TCCCATATCACCGCTCACCGTCTTTTCATTGCCATACGGAATTCCGGATGAGCATTTCATC
AGGCGGGCAAGAAATGTGAATAAAGGCCGATAAACTTGTGCTTATTTTCTTTACGGTC
TTTAAAAAGGCCGTAATATCCAGCTGAACGGTCTGGTTATAGGTACATTGAGCAACTGAC
TGAAATGCCTCAAAATGTTCTTTACGATGCCATTGGGATATATCAACGGTGGTATATCCA
GTGATTTTTTCTCCATTTTAGCTTCCTTAGCTCCTGAAAATCTCGATAACTCAAAAAAT
ACGCCGGTAGTGATCTTATTTCATTATGGTGAAAGTTGGAACCTCTTACGTGCCGATCA
ACGTCTCATTTTTCGCAAAAAGTTGGCCCAAGGGCTTCCCGGTATCAACAGGGACACAGGA
TTTATTTATTCTGCGAAGTGATCTTCCGTCACAGGTATTTATTTCGGCGCAAAGTGCGTCG
GGTGATGCTGCCAACTTAGTCGACTACAGGTCACCTAATACCATCTAAGTAGTTGATTTCAT
AGTGACTGGATATGTTGTGTTTACAGTATTATGTAGTCTGTTTTTTATGCAAAATCTAA
TTTAAATATATTGATATTTATATCATTTTACGTTTCTCGTTCAGCTTCTTGTACAAAGTT
GGCATTATAAGAAAGCATTGCTTATCAATTTGTTGCAACGAACAGGTCACCTATCAGTCAA
AATAAAATCATTATTTGCCATCCAGCTGCAGCTCTGGCCCGTGTCTCAAAATCTCTGATG
TTACATTGCACAAAGATAAAAAATATATCATCATGAATTCATGTTTGACAGCTTATCATC
GATAAGCTTTAATGCGGTAGTTTATCACAGTTAAATTGCTAACGCAGTCAGGCACCGTGT
ATGAAATCTAACAATGCGCTCATCGTCATCCTCGGCACCGTCACCTGGATGCTGTAGGC
ATAGGCTTGGTTATGCGGTAAGTCTGCGGGCTCTTGGCGGATATCGTCCATTCCGACAGC
ATCGCCAGTCACTATGGCGTGCTGTAGCGCTATATGCGTTGATGCAATTTCTATGCGCA
CCCGTTCTCGGAGCACTGTCCGACCGCTTGGCCGCGCCAGTCTGCTCGCTTCGCTA
CTTGAGGCCACTATCGACTACGCGATCATGGCGACCACACCGTCTGTTGGATCCTCTAC
GCCGAGCAGCATCGTGGCCGGCATCACCGGCGCCACAGGTGCGGTTGCTGGCGCCTATATC
GCCGACATCACCGATGGGGAAGATCGGGCTCGCCACTTCGGGCTCATGAGCGCTTGTTC
GGCGTGCGTATGGTGGCAGGCCCGTGGCCGGGGGACTGTTGGGCGCCATCTCCTTGCT
GCACCATTCCTTGGCGCGCGGTGCTCAACGGCCTCAACCTACTACTGGGCTGCTTCCTA
ATGCAAGGAGTCGCATAAGGGAGAGCGTCGACCGATGCCCTTGAGAGCCTTCAACCCAGTC
AGCTCCTTCCGGTGGGCGCGGGGATGACTATCGTCCCGCACCTTATGACTGCTTCTTT
ATCATGCAACTCGTAGGACAGGTGCCGGCAGCGCTCTGGGTCATTTTCGGCGAGGACCGC
TTTCGCTGGAGCGCGACGATGATCGGCCTGTGCTTGGCGGTATTCGGAATCTTGACGCC
CTCGCTCAAGCCTTCGTCAGTGGTCCCGCCACCAACGTTTCGGCGAGAAGCAGGCCATT
ATCGCCGGCATGGCGGCCGACGCGCTGGGCTACGCTTGGCTGGCGTTTCGCGACGCGAGGC
TGGATGGCCTTCCCCATTATGATTCTTCTCGCTTCCGGCGGCATCGGGATGCCCGCGTTG
CAGGCCATGCTGTCCAGGCAAGTAGATGACGACCATCAGGGACAGCTTCAAGGATCGCTC
GCGGCTCTTACAGCCTAATTCGATCATTGGACCGCTGATCGTCACGGCGATTATGCC
GCCTCGGCGAGCATGGAACGGGTTGGCATGGATTGTAGGCGCCGCCCTATACCTTGTG
TGCTTCCCGCGTTGCGTCCGCGGTGCTAGGAGCCGGGCCACCTCGACCTGAATGGAAGCC
GGCGGCACCTCGCTAACGGATTCAACACTCCAAGAATTGGAGCCAATCAATCTTTCGCGA
GAACTGTGAATGCGCAAAACCAACCTTGGCAGAACATATCCATCGCATGACCAAAATCCC
TTAAGCTGAGTTTTCTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTTC
TTGAGATCCTTTTTTCTGCGCGTAATCTGCTGCTTGCAAAACAAAAAACACCGCTACC
AGCGGTGGTTTTGTTTGGCGGATCAAGAGCTACCAACTCTTTTTCCGAAGGTAAGTGGCTT
CAGCAGAGCGCAGATACCAATACTGTCTTCTAGTGTAGCCGTAGTTAGGCCACCACTT
CAAGAACTCTGTAGCACCGCCTACATACCTCGCTCTGCTAATCCTGTTACAGTGGCTGC
TGCCAGTGGCGATAAGTCGTGCTTACCGGGTTGGACTCAAGACGATAGTTACCGGATAA
GGCGCAGCGTTCGGGCTGAACGGGGGTTTCGTGCACACAGCCAGCTTGGAGCGAACGAC
CTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCGAAGG
GAGAAAGGCGGACAGGTATCCGGTAAGCGGCAGGTCGGAACAGGAGAGCGCACGAGGGA
GCTTCCAGGGGGAACGCCTGGTATCTTTATAGTCTGTGCGGTTTCGCCACCTCTGACT
TGAGCGTCGATTTTTGTGATGCTCGTCAGGGGGCGGAGCCTATGGAACCGCCAGCAA-

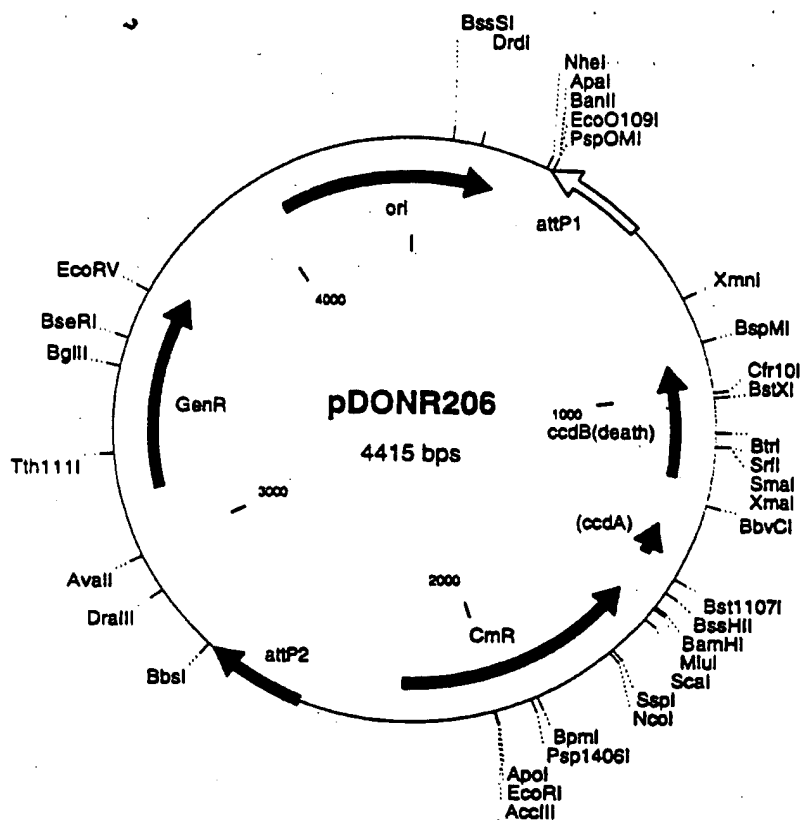
FIGURE 53B

158/240

CGCGGCCTTTTACGGTTCTTGGCCTTTTGCTGGCCTTTTGCTCACATGTTCTTTCCTGC
GTTATCCCCTGATTCTGTGGATAACCGTATTACCGCTAGCCAGGAAGAGTTTGTAGAAAC
GCAAAAAGGCCATCCGTCAGGATGGCCTTCTGCTTAGTTTGATGCCTGGCAGTTTATGGC
GGGCGTCCTGCCCCCACCCTCCGGGCCGTTGCTTCACAACGTTCAAATCCGCTCCCGGC
GGATTTGTCTACTCAGGAGAGCGTTCACCGACAAACAACAGATAAAACGAAAGGCCAG
TCTTCCGACTGAGCCTTTCGTTTTATTGATGCTGGCAGTTCCCTACTCTCGCGTTAAC
GCTAGCATGGATCTCGGGCCCCAAATAATGATTTTATTTTGACTGATAGTGACCTGTTG
TTGCAACAAATTGATGAGCAATGCTTTTTATAATGCCAATTTGTACAAAAAGCTGAA
CGAGAAACGTAAAATGATATAAATATCAATATATTAAATTAGATTTTGCATAAAAAACAG
ACTACATAATACTGTAAAACACAACATATCCAGTCACTATGAATCAACTACTTAGATGGT
ATTAGTGACCTGTAGTCGACCGACAGCCTTCCAAATGTTCTTCGGGTGATGCTGCCAAT
TAGTCGACCGACAGCCTTCCAAATGTTCTTCTCAAACGGAATCGTCGTATCCAGCCTACT
CGCTATTGTCCTCAATGCCGTATTAAATCATAAAAGAAATAAGAAAAAGAGGTGCGAGC
CTCTTTTTTGTGTGACAAAATAAAAACATCTACCTATTTCATATACGCTAGTGTCTAGTC
CTGAAAATCATCTGCATCAAGAACAATTTCACTACTTTTCTTCTTACAAGTCG
TTCGGCTTCATCTGGATTTTCAGCCTCTATACTTACTAAACGTGATAAAGTTTCTGTAAT
TTCTACTGTATCGACCTGCAGACTGGCTGTGTATAAGGGAGCCTGACATTATATTCCCC
AGAACATCAGGTTAATGGCGTTTTTGATGTCATTTTCGCGGTGGCTGAGATCAGCCACTT
CTTCCCCGATAACGGAGACCGGCACACTGGCCATATCGGTGGTCATCATGCGCCAGCTTT
CATCCCCGATATGCACCACCGGGTAAAGTTTACCGGGAGACTTATCTGACAGCAGACGTG
CACTGGCCAGGGGGATCACCATCCGTGCGCCGGGCGTGTCAATAATATCACTCTGTACAT
CCACAAACAGACGATAACGGCTCTCTCTTTTATAGGTGTAAACCTTAACTGCATTTTAC
CAGTCCCTGTTCTCGTCAGCAAAAGAGCCGTTCAATTAACCGGGCGACCTCAGCC
ATCCCTTCCCTGATTTTCCGCTTTCCAGCGTTCGGCACGCAGACGACGGGCTTCATTCTGC
ATGGTTGTGCTTACCAGACCGGAGATATTGACATCATATATGCCTTGAGCAACTGATAGC
TGTCGCTGTCAACTGTCACTGTAATACGCTGCTTCATAGCACACCTCTTTTGTACATACT
TCGGGTATACATATCAGTATATATTCTTATACCGCAAAATCAGCGCGCAAAATACGCATA
CTGTTATCTGGCTTTTAGTAAGCCGGATCCACGCGATTACGCCCCGCCCTGCCACTCATC
GCAGTACTGTTGTAATTCATTAAGCATTCGCGGACATGGAAGCCATCACAGACGGCATG
ATGAACCTGAATCGCCAGC

FIGURE 53C

159/240



pDONR206 4415 bp

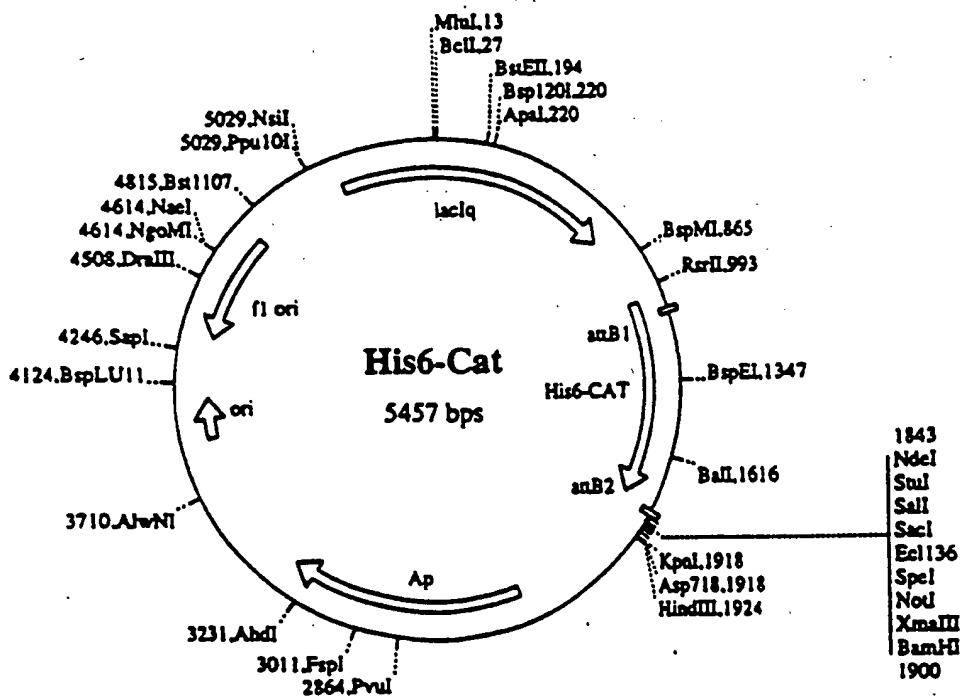
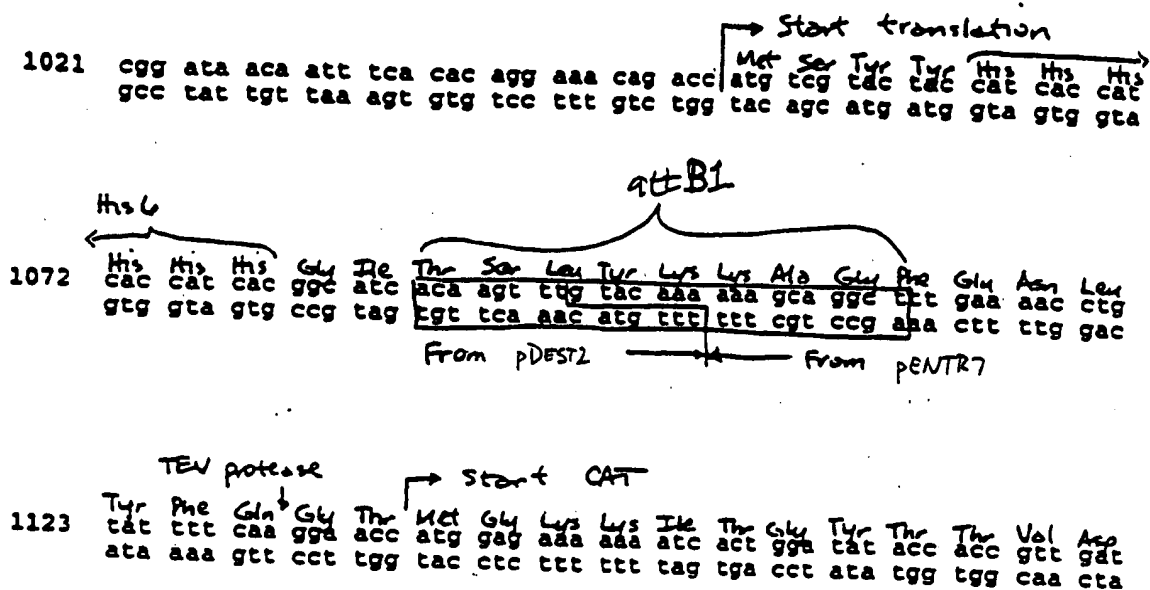
CGGCATTGAGGACAATAGCGAGTAGGCTGGATACGACGATTCCGTTTGAGAAGAACATTT
GGAAGGCTGTCCGGTCGACTACAGGTCACTAATACCATCTAAGTAGTTGAATCATAGTGAC
TGGATATGTTGTGTTTTACAGTATTATGTAGTCTGTTTTTATGCAAAATCTAATTTAAT
ATATTGATATTTATATCATTTTTACGTTTTCTCGTTCAGCTTTTTTGTACAAAGTTGGCATT
ATAAAAAAGCATTGCTTATCAATTTGTTGCAACGAACAGGTCACTATCAGTCAAAATAAA
ATCATTATTTGGGGCCCGAGATCCATGCTAGCGGTAATACGGTTATCCACAGAATCAGGG
GATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAG
GCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCAGAAAAATCGA
CGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACAGGCGTTTCCCCCT
GGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCC
TTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCCG
GTGTAGGTCTGTTCTGCTCCAAAGCTGGGCTGTGTGCACGAACCCCCGTTTCCGCCCCGCGC
TGCGCCTTATCCGGTAACATATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCA
CTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAG
TTCTTGAAGTGGTGGCCTAACTACGCTACACTAGAAGGACAGTATTTGGTATCTGCGCT
CTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAAACAAAC
ACCGCTGGTAGCGGTGGTTTTTTTGTGTTGCAAGCAGCAGATTACGCGCAGAAAAAAGGA
TCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGAACGAAAACTCA
CGTTAAGGGATTTTGGTCATGNCGCCCTCCGCTCAAGTCAGCGTAATGCTCTGCCAGTGT
TACAACCAATTAACCAATTCTGATTAGAAAACTCATCGAGCATCAAATGAACTGCAAT
TTATTATATCAGGATTATCAATACCATATTTTTGAAAAAGCGTTTCTGTAATGAAGGA
GAAAACTCACCGAGGCGAGTTCATAGGATGGCAAGATCCTGGTATCGGTCTGCGATTCCG
ACTCGTCCAACATCAATACAACCTATTAGCCGAGGTCTTCCGATCTCCTGAAGCCAGGGC
AGATCCGTGCACAGCACCTTGCCGTAGAAGAACAGCAAGGCCGCCAATGCCTGACGATGC
GTGGAGCCGAAACCTTGCGCTCGTTGCCAGCCAGGACAGAAATGCCTCGACTTCCGCTG
CTGCCCAAGGTTGCCGGGTGACGACACCCGTGGAAACGGATGAAGGCACGAACCCAGTTG
ACATAAGCCTGTTCCGTTCTGTAACTGTAATGCAAGTAGCGTATGCGCTCACGCAACTGG
TCCAGAACCCTTGACCGAACGCAGCGGTGGTAACGGCGCAGTGGCGGTTTTTCATGGCTTGT
TATGACTGTTTTTTTGTACAGTCTATGCCTCGGGCATCCAAGCAGCAAGCGCGTTACGCC
GTGGGTTCGATGTTTGTATGTTATGGAGCAGCAACGATGTTACGCAGCAGCAACGATGTTAC
GCAGCAGGGCAGTCGCCCTAAAAACAAAGTTAGGTGGCTCAAGTATGGGCATCATTCGCAC
ATGTAGGCTCGGCCCTGACCAAGTCAAATCCATGCGGGCTGCTCTTGATCTTTTCCGGTCG
TGAGTTCCGGAGAGCTAGCCACCTACTCCCAACATCAGCCGGAATCCGATTACCTCGGGAA
CTTGCTCCGTAGTAAGACATTATCGCGCTTGCTGCTTCGACCAAGAAGCGGTTGTTGG
CGCTCTCGCGGCTTACGTTCTGCCCAGGTTTGAGCAGCCGCTAGTGAGATCTATATCTA
TGATCTCGCAGTCTCCGGCAGCACCGGAGGCAGGGCATTGCCACCGCGCTCATCAATCT
CCTCAAGCATGAGGCCAACGCGCTTGGTCTTATGTGATCTACGTGCAAGCAGATTACGG
TGACGATCCCGCAGTGGCTCTCTATACAAAGTTGGGCATACGGGAAGAAGTGATGCACTT
TGATATCGACCCAAGTACCGCCACCTAACAATTCGTTCAAGCCGAGATCGGCTTCCCGGC
CTAATTTCCCTCGTCAAAAAAAGGTTATCAAGTGAGAAATCACCATGAGTGACGACTG
AATCCGGTGAGAATGGCAAAAGCGTATGCATTTCTTCCAGACTTGTTCACAGGCCAGC
CATTACGCTCGTCATCAAAATCACTCGCATCAACCAACCGTTATTTCATTCTGTGATTGCG
CCTGAGCGAGACGAAATACGCGATCGCTGTTAAAGGACAATTACAAACAGGAATCGAAT
GCAACCGGCGCAGGAACACTGCCAGCGCATCAACAATATTTTACCTGAATCAGGATATT
CTTCTAATACCTGGAATGCTGTTTTCCCGCGGATCGCAGTGGTGAAGTAAACATGCATCAT
CAGGAGTACGGATAAAATGCTTGATGGTCCGAAGAGGCATAAATCCGTCAGCCAGTTTA
GTCTGACCATCTCATCTGTAACATCATTGGCAACGCTACCTTTGCCATGTTTTAGAAACA
ACTCTGGCGCATCGGGCTTCCCATACAATCGAAAGATTGTGCGACCTGATTGCCCGACAT
TATCGCGAGCCCATTTATACCCATATAAATCAGCATCCATGTTGGAATTTAATCGCGGCC
TCCAGCAAGACGTTTCCCGTTGAATATGGCTCATAACACCCCTTGTATTACTGTTTATGT
AAGCAGACAGTTTTATTGTTTCATGATGATATATTTTTATCTGTGCAATGTAACATCAGA
GATTTTGAGACACGGGCCNCGCGCACTGCAGCTGGATCGGCAATAATGATTTTTATTTTG
ACTGATAGTGACCTGTTCTGTTGCAACAAATTGATAAGCAATGCTTTTTTATAATGCCAAC -

161/240

TTTGTACAAGAAAGCTGAACGAGAAACGTAAAATGATATAAATATCAATATATTAAATTA
GATTTTGCATAAAAAACAGACTACATAACTGTAAAACACAACATATCCAGTCACTATG
ATTCAACTACTTAGATGGTATTAGTGACCTGTAGTCGACTAAGTTGGCAGCATCACCCGA
CGCACTTTGCGCCGAATAAATACCTGTGACGGAAGATCACTTCGCAGAATAAATAAATCC
TGGTGTCCCTGTTGATACCGGGAAGCCCTGGGCCAACTTTTGGCGAAAATGAGACGTTGA
TCGGCAGCTAAGAGGTTCCAACCTTTCACCATAATGAAATAAGATCACTACCGGGCGTATT
TTTTGAGTTATCGAGATTTTCAGGAGCTAAGGAAGCTAAAATGGAGAAAAAATCACTGG
ATATACCACCGTTGATATATCCCAATGGCATCGTAAAGAACATTTTGAGGCATTTCAGTC
AGTTGCTCAATGTACCTATAACCAGACCGTTCAGCTGGATATTACGGCCTTTTAAAGAC
CGTAAAGAAAAATAAGCACAAAGTTTTATCCGGCCTTTATTACATTCTTGCCCGCCTGAT
GAATGCTCATCCGGAATTCCGTATGGCAATGAAAGACGGTGAGCTGGTGATATGGGATAG
TGTTCAACCTTGTACACCGTTTTCCATGAGCAAACCTGAAACGTTTTTCATCGCTCTGGAG
TGAATACCACGACGATTTCCGGCAGTTTCTACACATATATTGCAAGATGTGGCGTGTTA
CGGTGAAAACCTGGCCTATTTCCCTAAAGGGTTTATTGAGAATATGTTTTTCGTCTCAGC
CAATCCCTGGGTGAGTTTCCACGATTTTGATTTAAACGTGGCCAATATGGACAACCTTCTT
CGCCCCCGTTTTCCCATGGGCAAATATTATACGCAAGGCGACAAGGTGCTGATGCCGCT
GGCGATTCCAGGTTTCATCATGCCGTCTGTGATGGCTTCCATGTCCGCAGAATGCTTAATGA
ATTACAACAGTACTGCGATGAGTGGCAGGGCGGGGCGTAAACGCGTGATCCGGCTTACT
AAAAGCCAGATAACAGTATGCGTATTTGCGCGCTGATTTTTGCGGTATAAGAATATATAC
TGATATGTATACCCGAAGTATGTCAAAAAGAGGTGTGCTATGAAGCAGCGTATTACAGTG
ACAGTTGACAGCGACAGCTATCAGTTGCTCAAGGCATATATGATGTCAATATCTCCGGTC
TGGTAAGCACAAACCATGCAGAATGAAGCCCGTCGTCTGCGTGCCGAACGCTGGAAAGCGG
AAAATCAGGAAGGGATGGCTGAGGTGCGCCCGTTTTATTGAAATGAACGGCTCTTTTGCTG
ACGAGAACAGGGACTGGTGAAATGCAGTTTAAGGTTTACACCTATAAAAGAGAGAGCCGT
TATCGTCTGTTTGTGGATGTACAGAGTGATATTATTGACACGCCCCGGGCGACGGATGGTG
ATCCCCCTGGCCAGTGACGCTCTGCTGTCAGATAAAGTCTCCCGTGAACTTTACCCGGTG
GTGCATATCGGGGATGAAAGCTGGCGCATGATGACCACCGATATGGCCAGTGTCGGGTC
TCCGTTATCGGGGAAGAAGTGGCTGATCTCAGCCACCGCGAAAATGACATCAAAAACGCC
ATTAACCTGATGTTCTGGGGAATATAAATGTCAAGGCTCCGTTATACACAGCCAGTCTGCA
GGTCGATACAGTAGAAATTACAGAACTTTATCACGTTTAGTAAGTATAGAGGCTGAAAA
TCCAGATGAAGCCGAACGACTTGTAAGAGAAAAGTATAAGAGTTGTGAAATGTTCTTGA
TGCAGATGATTTTCAGGACTATGACACTAGCATATATGAATAGGTAGATGTTTTATTTT
GTCACACAAAAAGAGGCTCGCACCTCTTTTTCTTATTTCTTTTATGATTTAATA

FIGURE 54 C

Figure 55 An Entry (pENTR7) Clone of CAT Subcloned into pDEST2



163/240

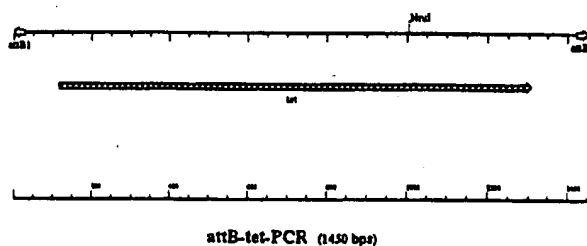
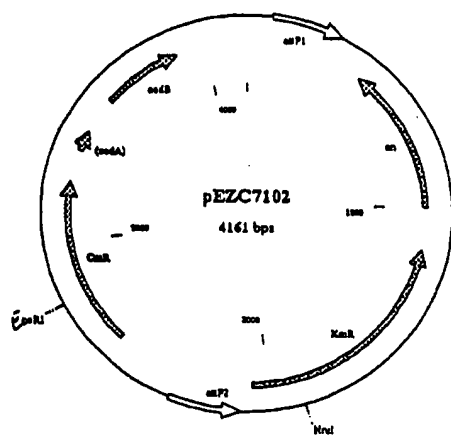


FIGURE 5b

164/240

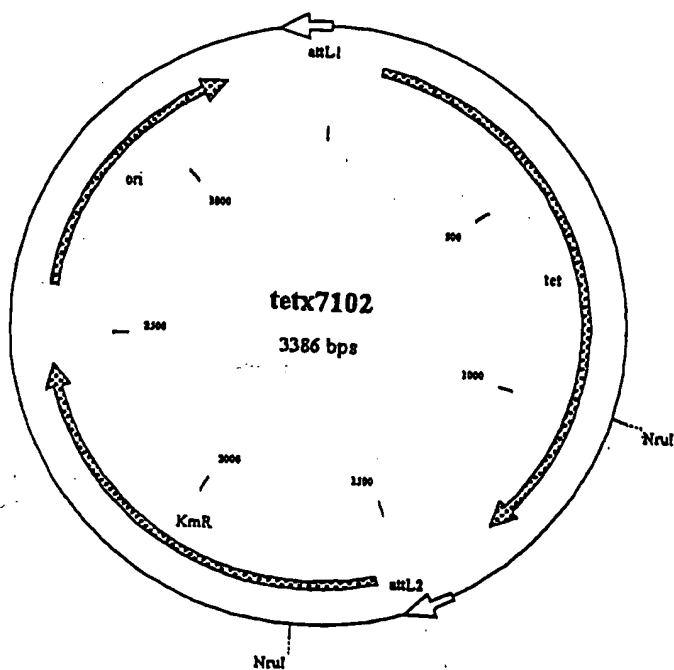


FIGURE 57

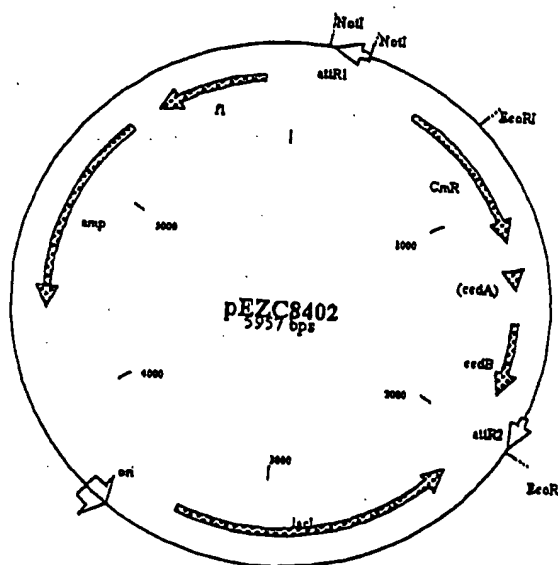


FIGURE 5B

166/240

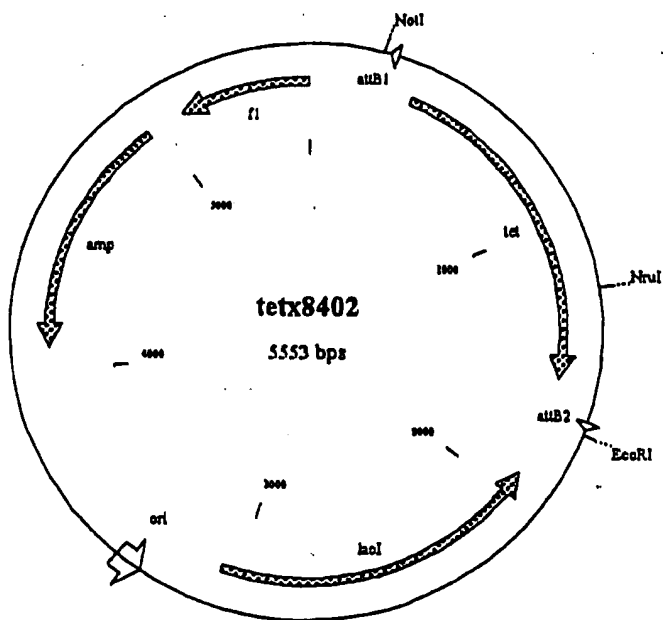


FIGURE 59

167/240

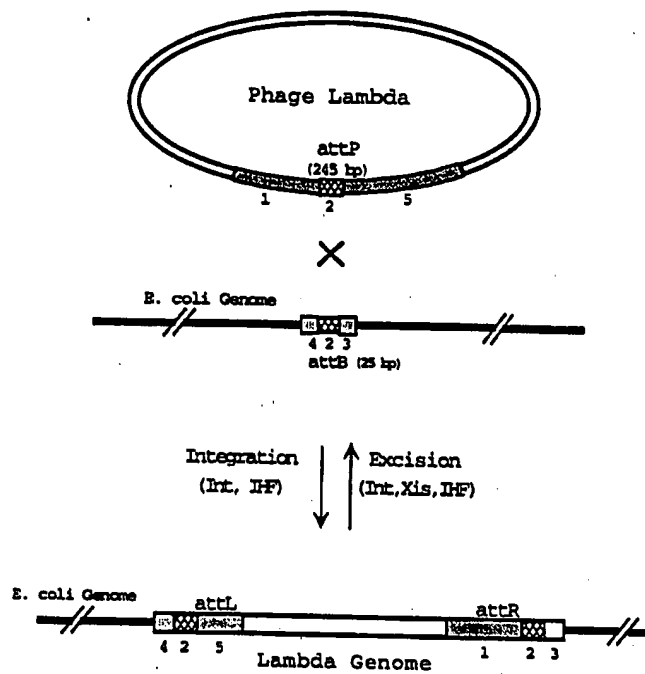


FIGURE 60

168/240

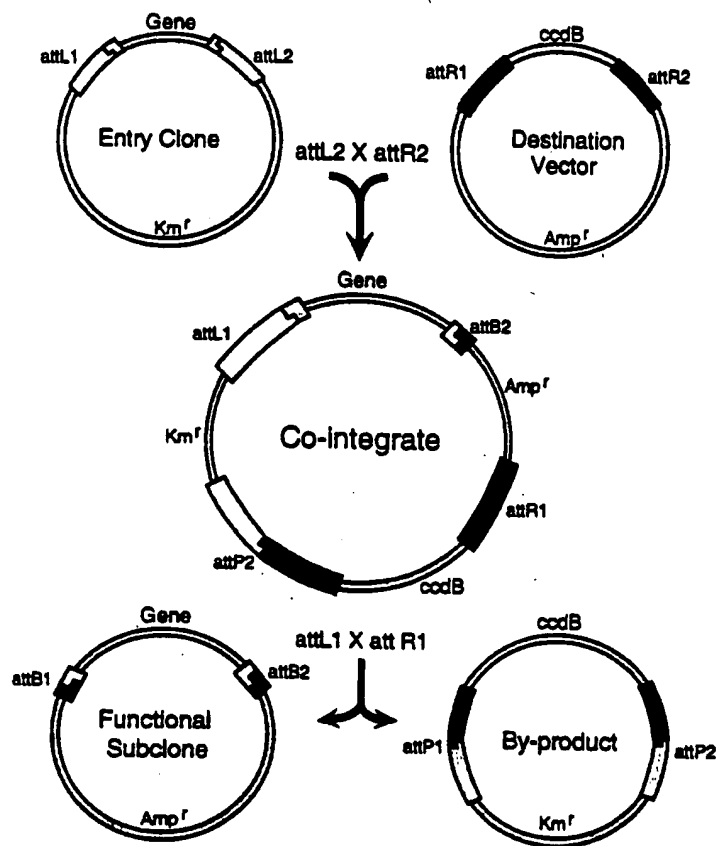


FIGURE 61

169/240

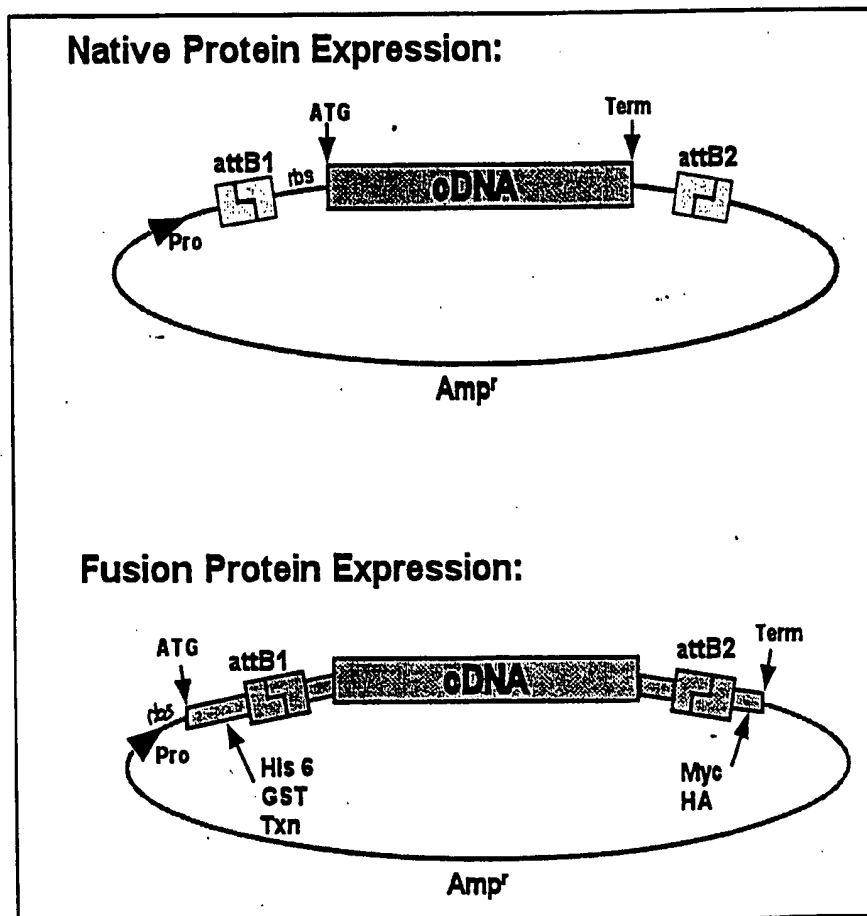


FIGURE 62

170/240

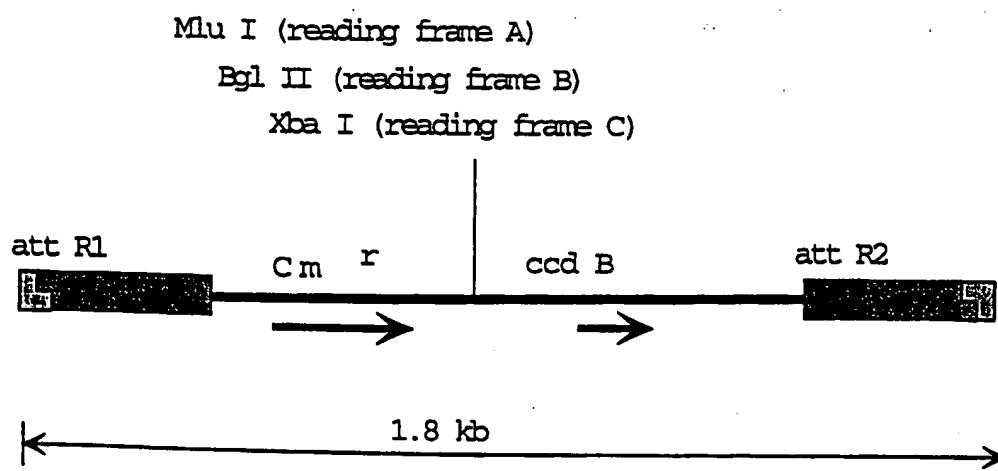
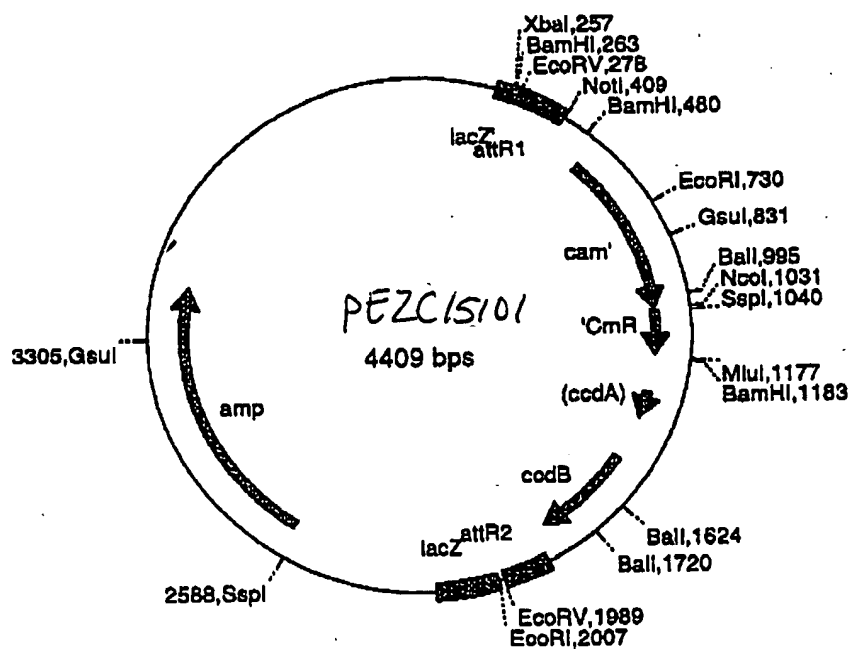


FIGURE 63

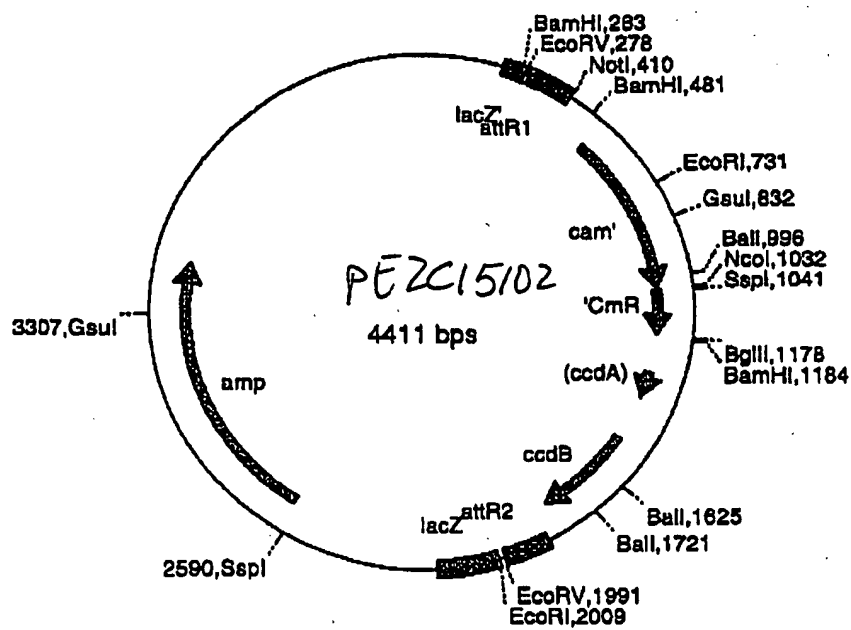
171/240

FIGURE 64A



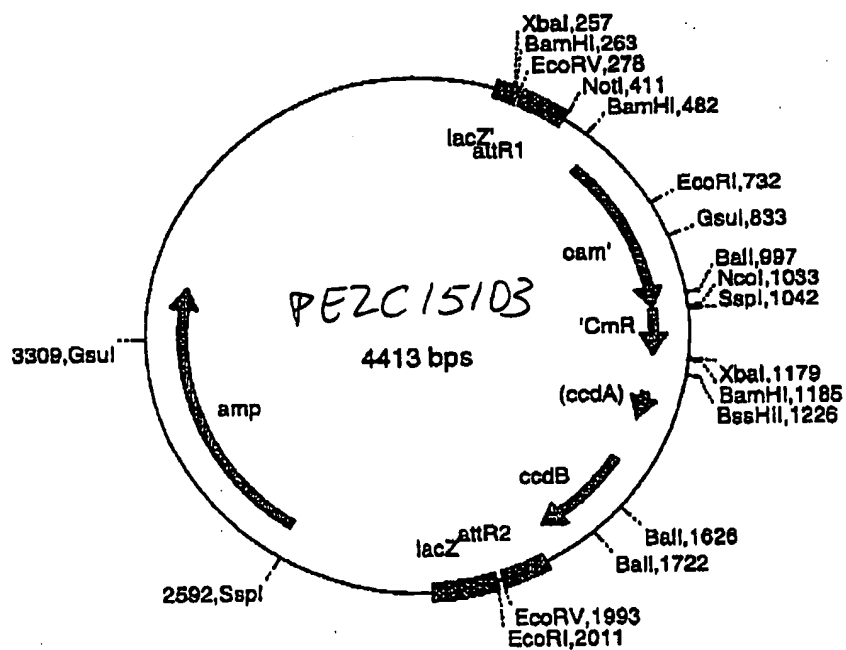
172/240

FIGURE 4A



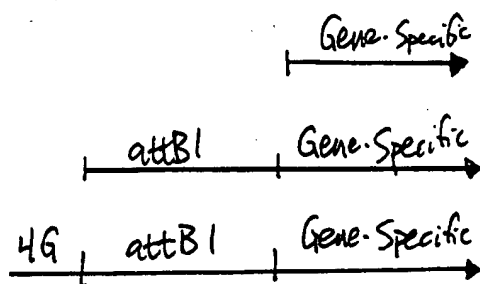
173/240

FIGURE 64C



Primers for Amplifying *tetR* and *ampR* for Cloning by Recombination

Primers



Reverse Primers

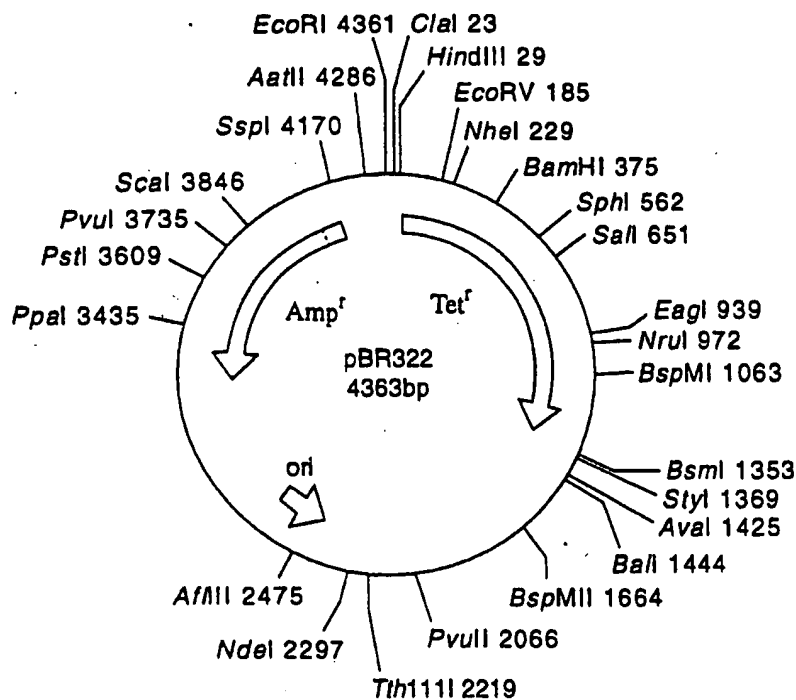
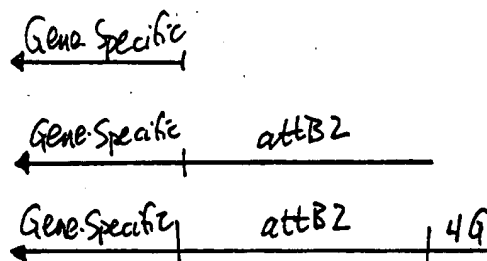


FIGURE 05

175/240

**Results of Cloning
tet and amp PCR Products
by Recombination**

PCR Product Used in GCS Reactions	No. Colonies Obtained (100 ul plated)	Form of DNA Analyzed	Colonies Obtained of Predicted Size
tet	6, 10	SC	0 of 8
attB-tet	9, 6	SC	1 of 8
attB+4G-tet	824, 1064	SC AvaI+Bam	7 of 7 7 of 7
amp	7, 13	SC	0 of 8
attB-amp	18, 22	SC	3 of 8
attB+4G-amp	3020, 3540	SC PstI	8 of 8 8 of 8
attB Plasmid (Pos. Control)	320, 394		

FIGURE 66

176/260

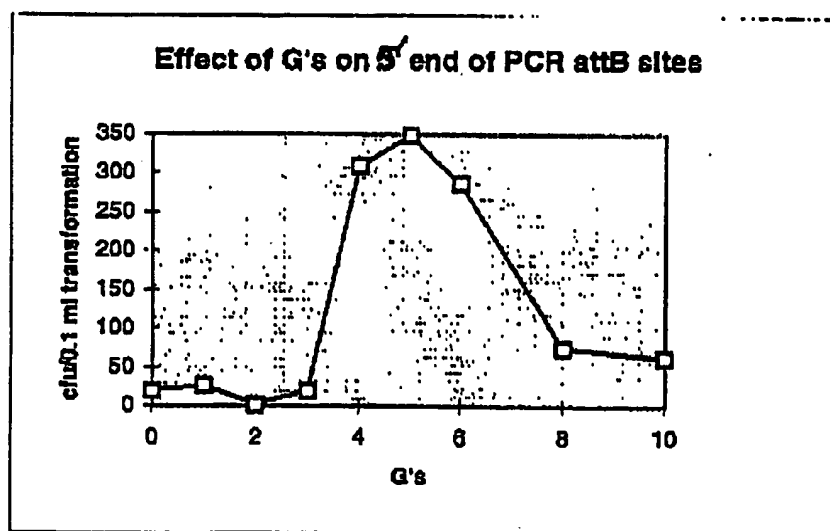


FIGURE 67

177/240

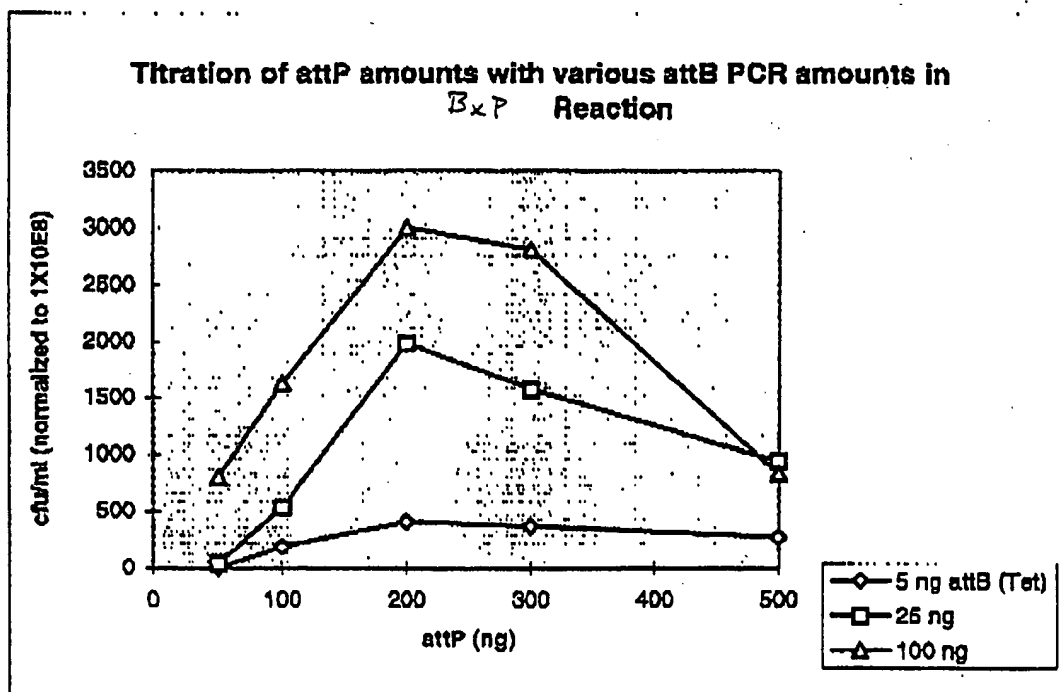
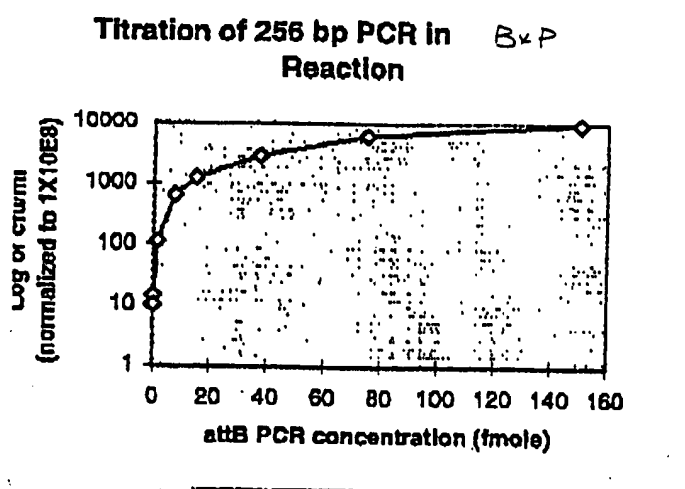


FIGURE 68

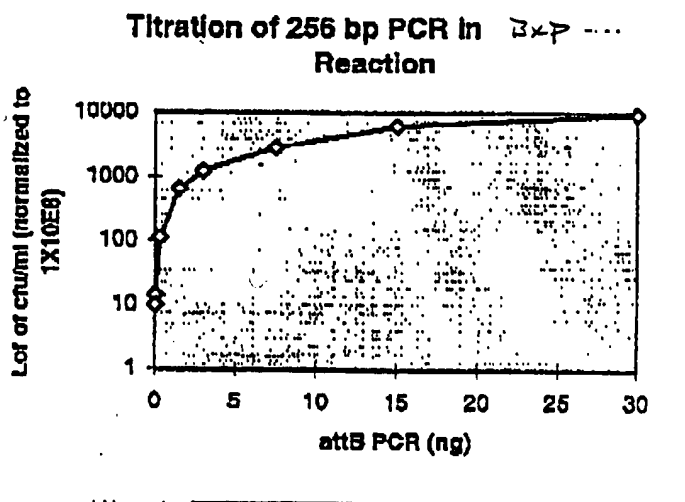
178/240

FIGURE
69

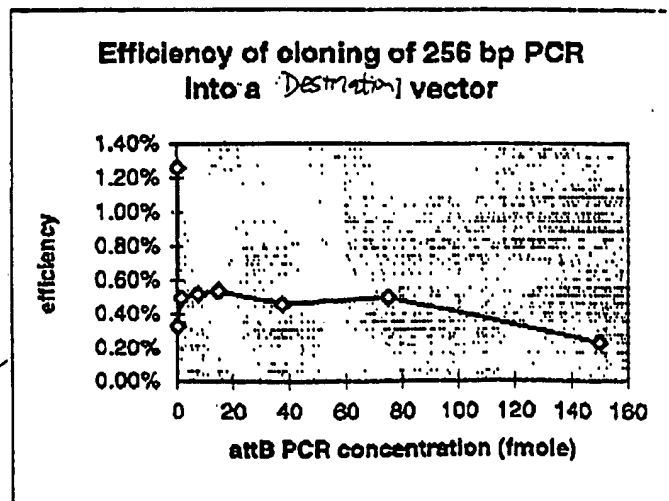
A



B



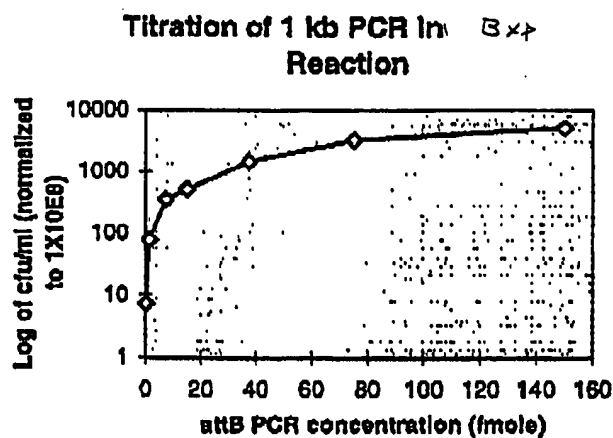
C



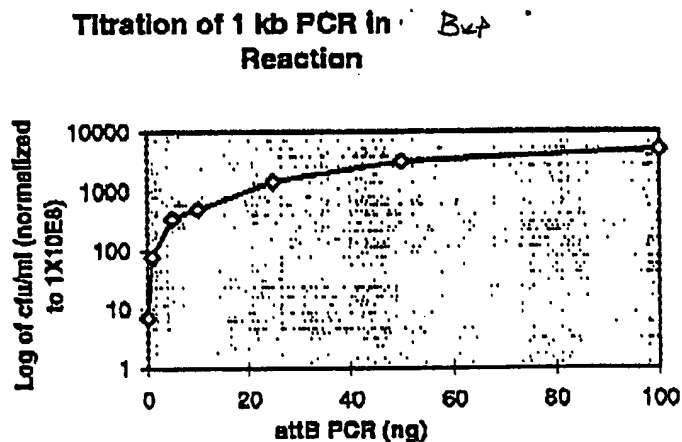
179/240

FIGURE
70

A



B



C

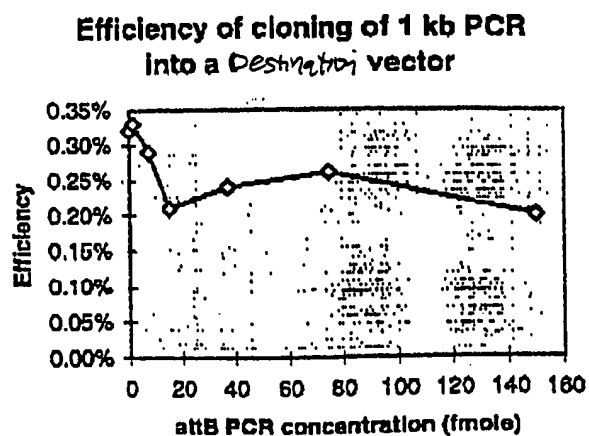
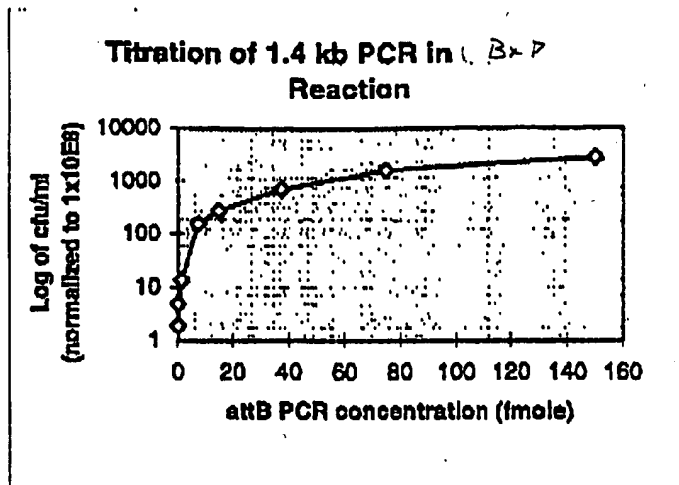
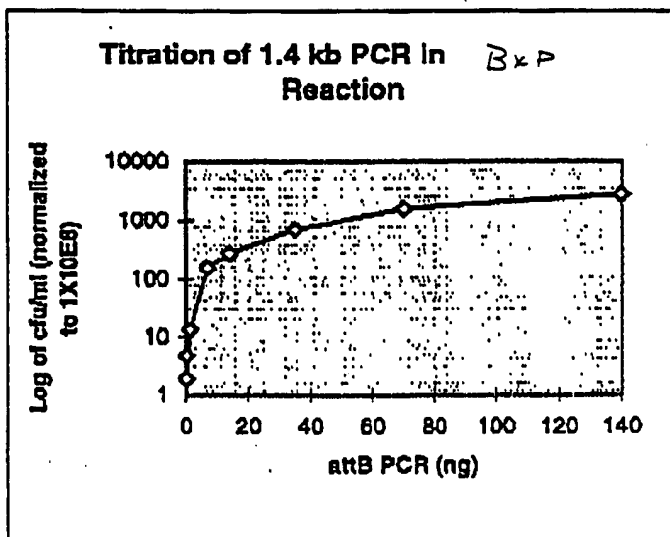


FIGURE 71

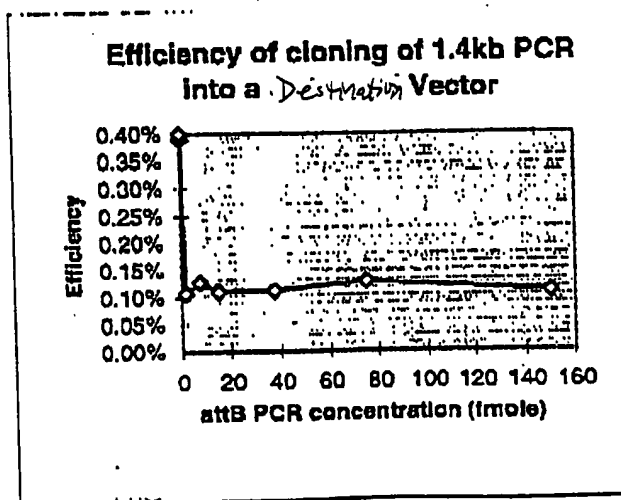
A



B



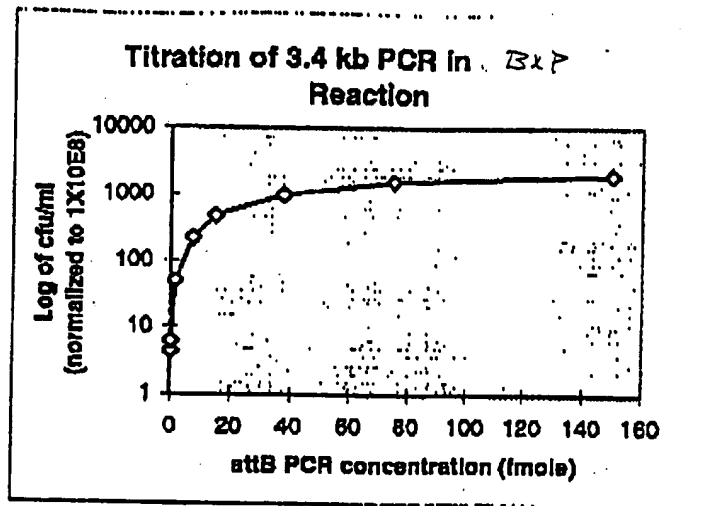
C



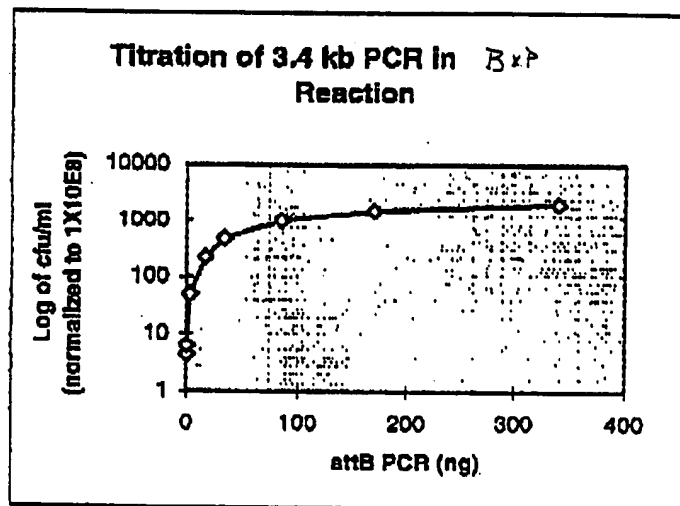
181/240

FIGURE 72

A



B



C

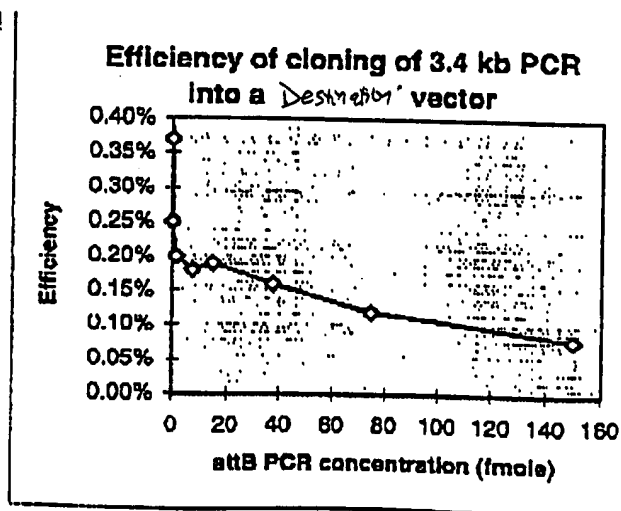
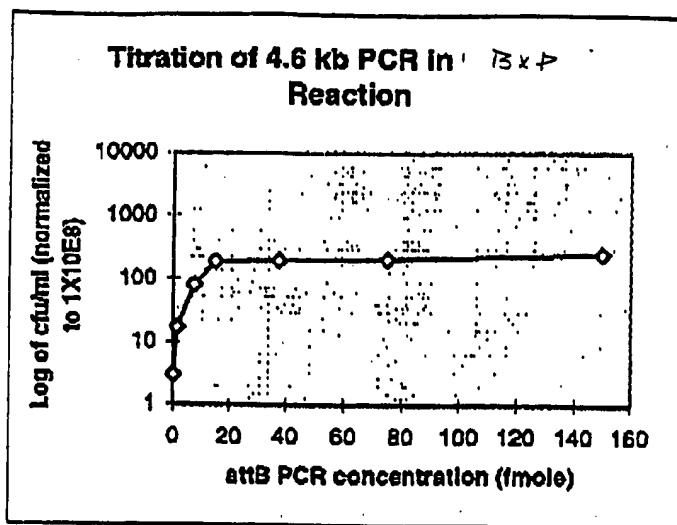
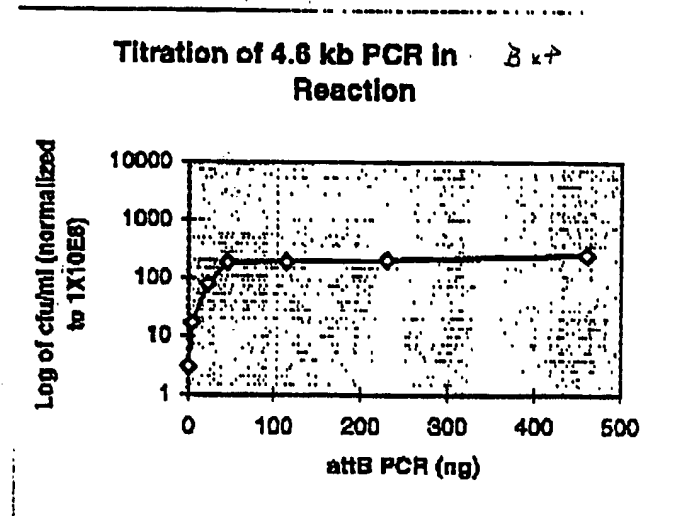


FIGURE 73

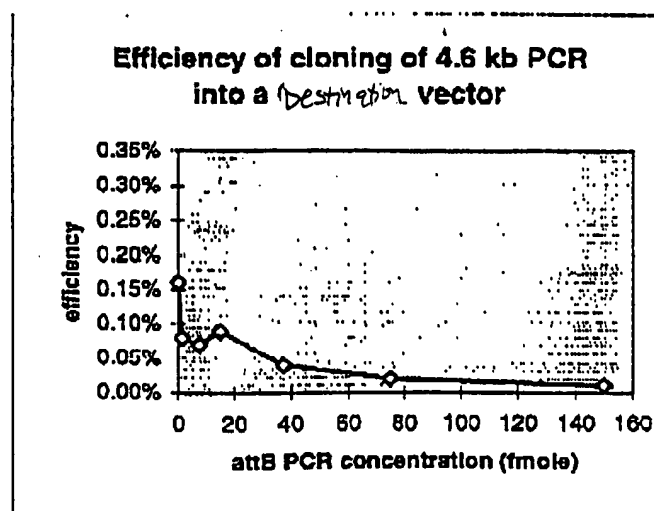
A



B



C



6.9 kb PCR DNA Titration In α B x P Reaction

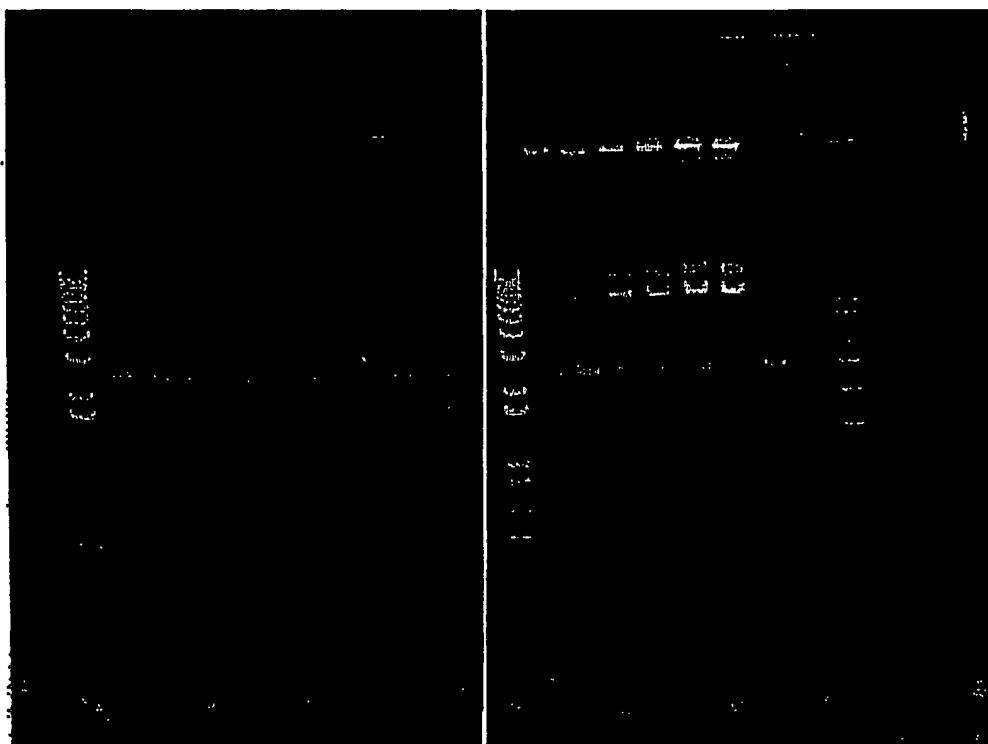


FIGURE 74

184/240

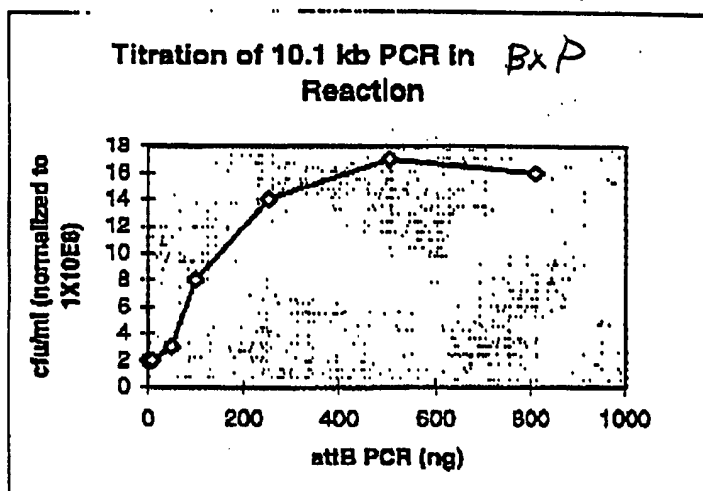


FIGURE 75-

185/240

10.1 kb PCR DNA Titration in Bx7 Reaction

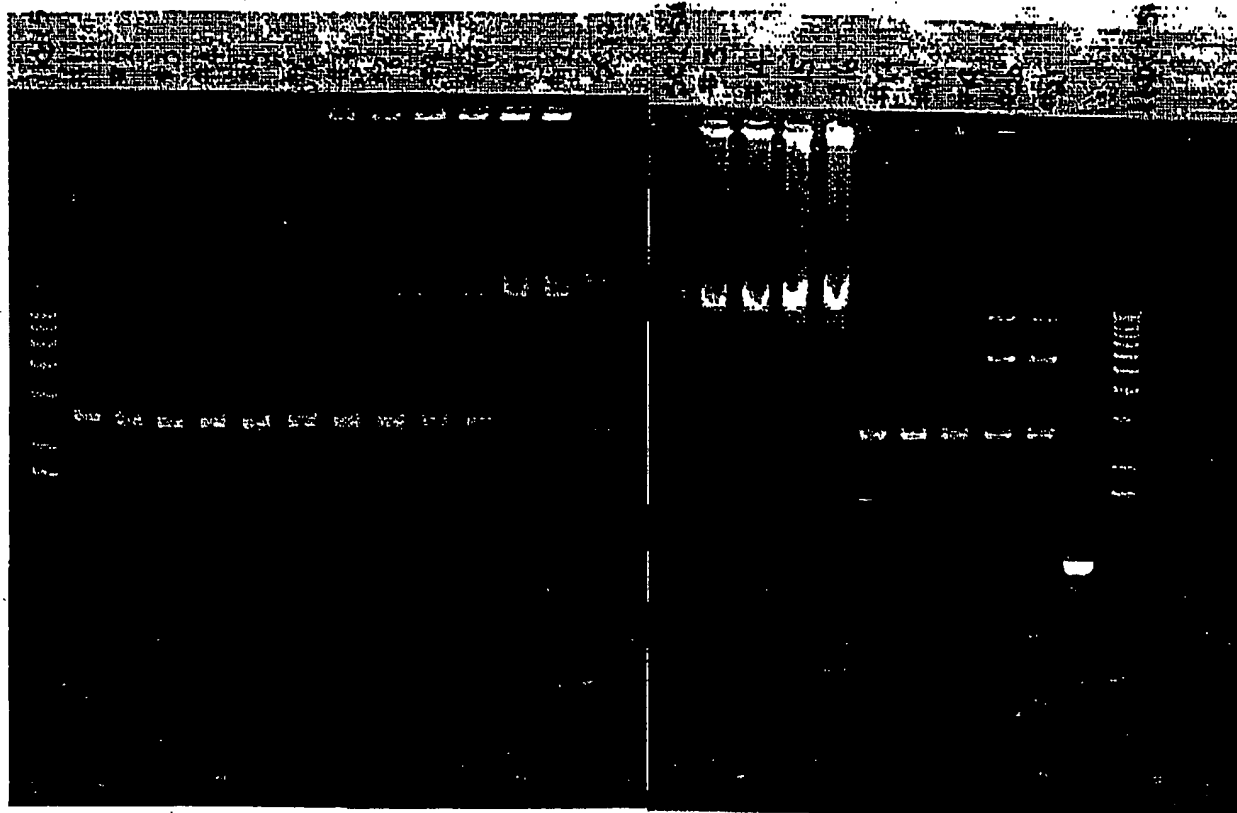


FIGURE 76

Cloning of PCR Products of Different Sizes with the GATEWAY™ PCR Cloning System

Size	fmols PCR DNA	ng PCR DNA	Cols/ml Transformation (pUC=10 ⁸ CFU/ml)	Correct Clones/Total Examined**
0.26 kb*	15	3	1223	10/10 (a)
	37.5	7.5	2815	
1.0 kb	15	10	507	49/50 (b)
	37.5	25	1447	
1.4 kb	15	14	271	48/50 (c)
	37.5	35	683	
3.4 kb	15	34	478	9/10 (a)
	37.5	85	976	
4.6 kb	15	46	190	10/10 (a)
	37.5	115	195	
6.9 kb	15	69	30 (235)**	47/50 (b)
	37.5	173	54 (463)**	

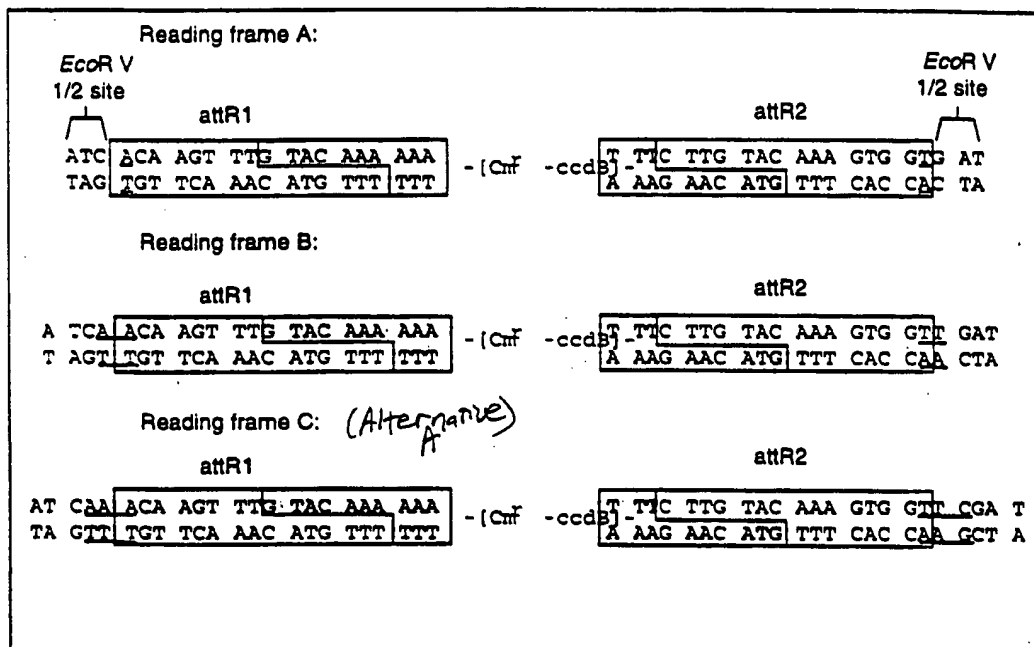
*The 0.26 kb PCR product was used unpurified; all the others were purified by precipitation with PEG/MgCl₂ as described in the text of Example 9, to remove primer dimers potentially present. Standard incubations were for 60 min.

**overnight incubation

- (a) DNA minipreps
- (b) ampR/kanR
- (c) tetR/kanR

Figure 77

187/240



Reading frame C: (Alternative)
B

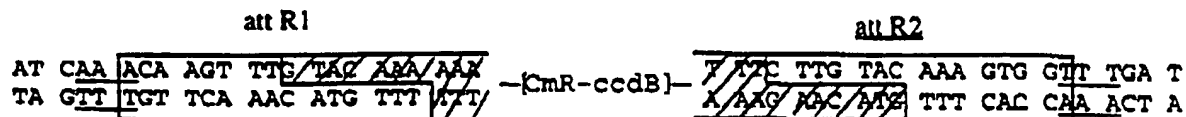


FIGURE 78

188/240

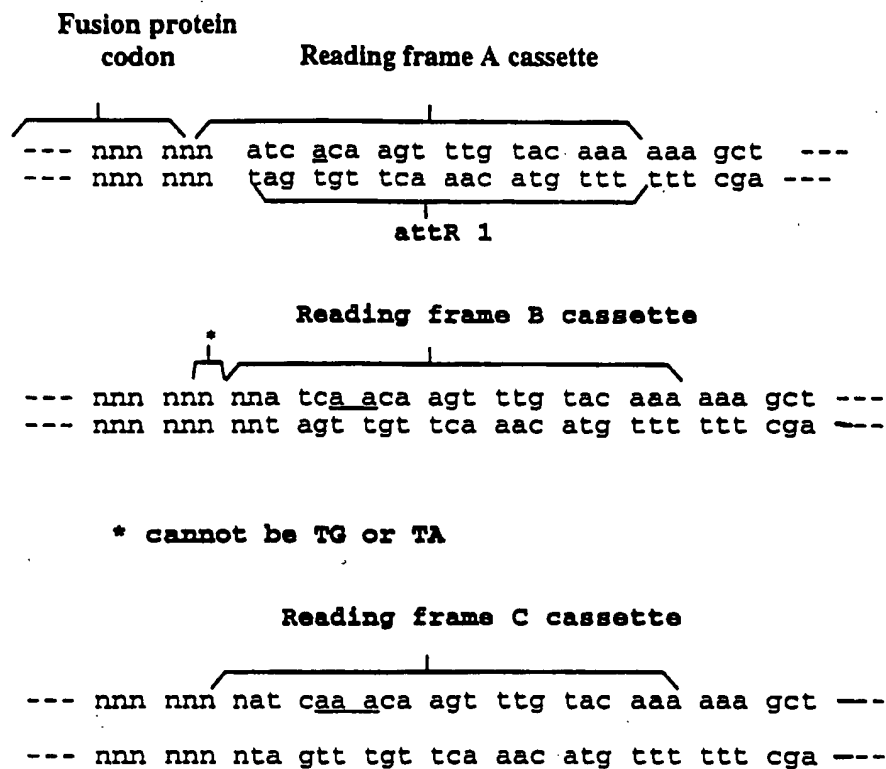


FIGURE 79

189/240

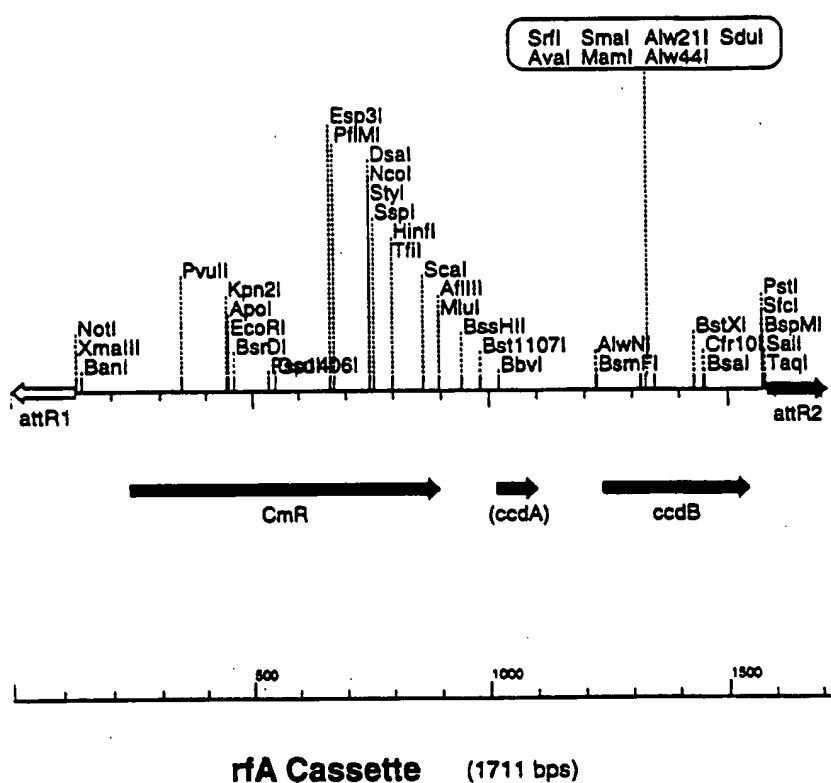


FIGURE 80

190/240

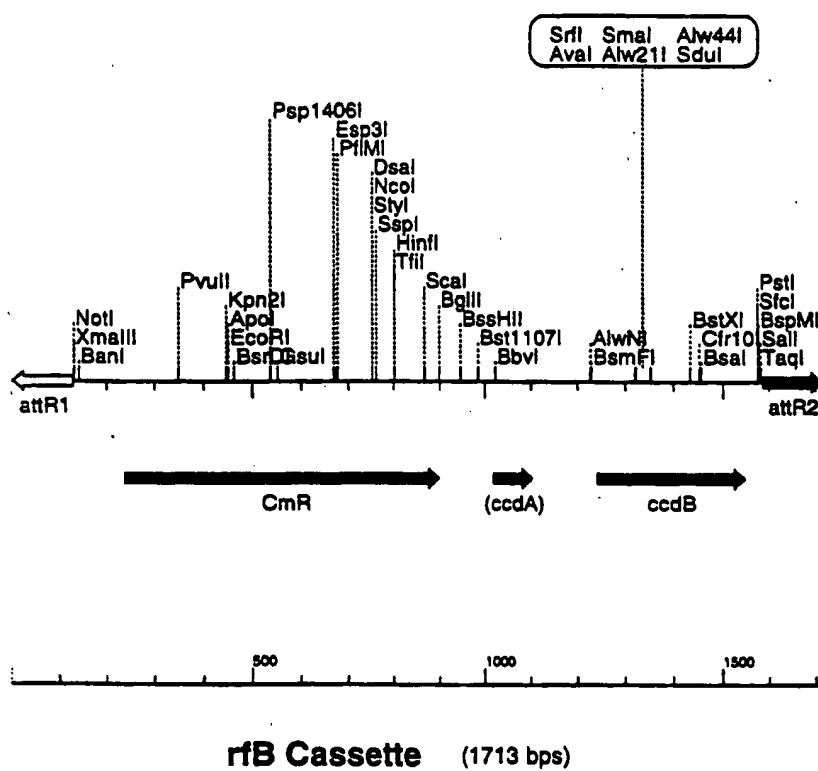
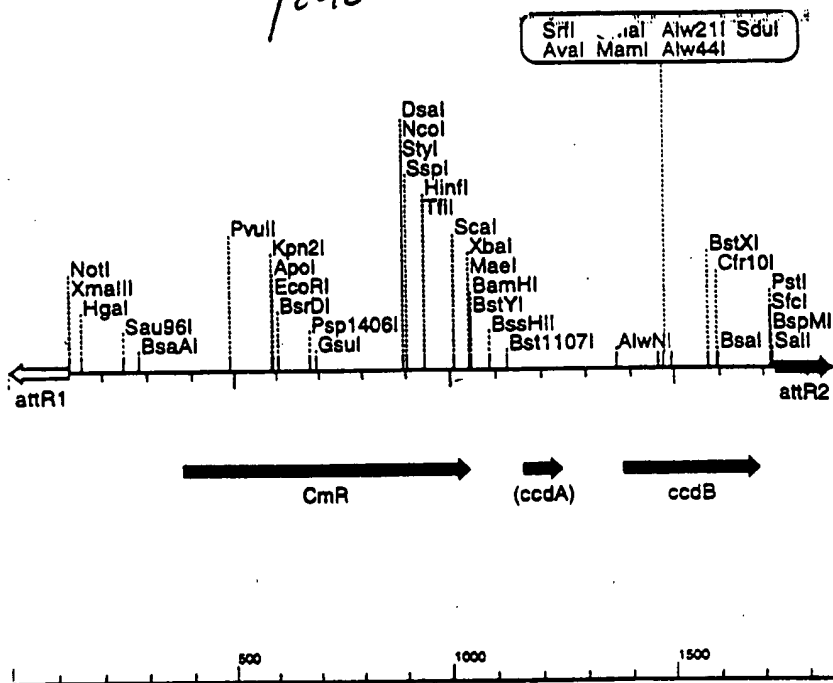


FIGURE 81

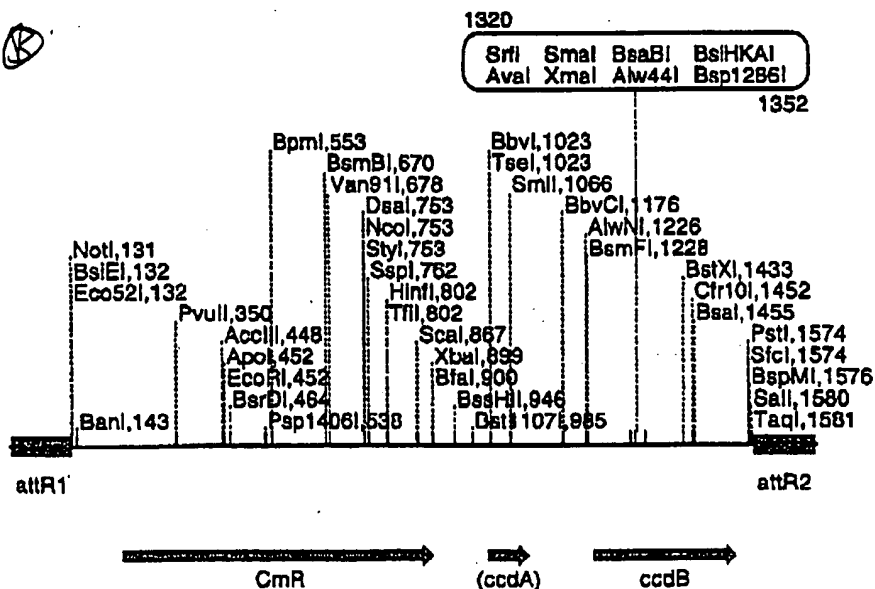
191/240

(A)



rfc Cassette (1856 bps)

(B)



rfc cassette (1715 bps)

FIGURE 82

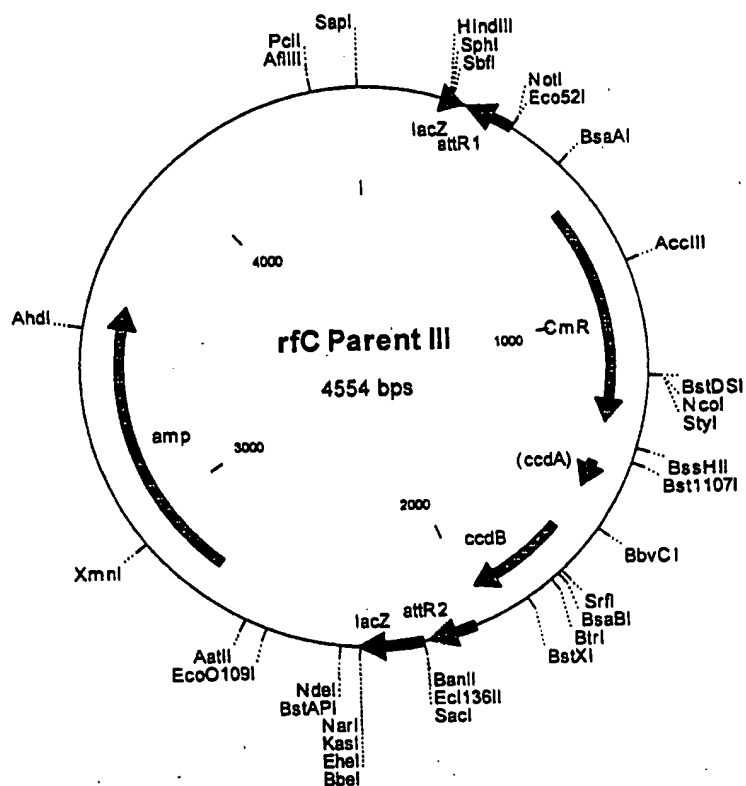


FIGURE 83 A

prfC Parent III 4554 bp

<u>Location (Base Nos.)</u>		<u>Gene Encoded</u>
410..286		attR1
660..1319		CmR
1439..1523		inactivated ccdA
1661..1966		ccdB
2007..2131		attR2
2753..3613		amp
1	GCGCCCAATA CGCAAACCGC CTCTCCCCGC GCGTTGGCCG ATTCATTAAT GCAGCTGGCA	
61	CGACAGGTTT CCCGACTGGA AAGCGGGCAG TGAGCGCAAC GCAATTAATG TGAGTTAGCT	
121	CACCTATTAG GCACCCACAG CTTTACACTT TATGCTTCCG GCTCGTATGT TGTGTGGAAT	
181	TGTGAGCGGA TAACAATTTC ACACAGGAAA CAGCTATGAC CATGATTACG CCAAGCTTGC	
241	ATGCCTGCAG GTCGACTCTA GAGGATCCCC GGGTACCGAT ATCAAACAAG TTTGTACAAA	
301	AAAGCTGAAC GAGAAACGTA AAATGATATA AATATCAATA TATTAAATTA GATTTTGCAT	
361	AAAAACAGA CTACATAATA CTGTAAAACA CAACATATCC AGTCACTATG GCGGCCGCTA	
421	AGTTGGCAGC ATCACCCGAC GCACTTTGCG CCGAATAAAT ACCTGTGACG GAAGATCACT	
481	TCGCAGAATA AATAAATCCT GGTGTCCCTG TTGATACCGG GAAGCCCTGG GCCAACTTTT	
541	GGCGAAAATG AGACGTTGAT CGGCACGTAA GAGGTTCCAA CTTTCACCAT AATGAAATAA	
601	GATCACTACC GGGCGTATTT TTTGAGTTAT CGAGATTTTC AGGAGCTAAG GAAGCTAAAA	
661	TGGAGAAAAA AATCACTGGA TATACCACCG TTGATATATC CCAATGGCAT CGTAAAGAAC	
721	ATTTTGAGGC ATTTCACTCA GTTGCTCAAT GTACCTATAA CCAGACCGTT CAGCTGGATA	
781	TTACGGCCTT TTTAAAGACC GTAAAGAAAA ATAAGCACA GTTTTATCCG GCCTTTATTC	
841	ACATTCTTGC CCGCCTGATG AATGCTCATC CGGAATTCCG TATGGCAATG AAAGACGGTG	
901	AGCTGGTGAT ATGGGATAGT GTTCACCCCT GTTACACCGT TTTCCATGAG CAAACTGAAA	
961	CGTTTTCAATC GCTCTGGAGT GAATACCACG ACGATTTCCG GCAGTTTCTA CACATATATT	
1021	CGCAAGATGT GGCCTGTTC GGTGAAAACC TGGCCTATTT CCCTAAAGGG TTTATTGAGA	
1081	ATATGTTTTT CGTCTCAGCC AATCCCTGGG TGAGTTTCAC CAGTTTGTAT TTAACGTTGG	
1141	CCAATATGGA CAACTTCTTC GCCCCCGTTT TCACCATGGG CAAATATTAT ACGCAAGGCG	
1201	ACAAGGTGCT GATGCCGCTG GCGATTCAAG TTCATCATGC CGTCTGTGAT GGCTTCCATG	
1261	TCGGCAGAAT GCTTAATGAA TTACAACAGT ACTGCGATGA GTGGCAGGGC GGGGCGTAAT	
1321	CTAGAGGATC CGGCTTACTA AAAGCCAGAT AACAGTATGC GTATTGTGCG GCTGATTTTT	
1381	GCGGTATAAG AATATATACT GATATGTATA CCCGAAGTAT GTCAAAAAGA GGTGTGCTAT	
1441	GAAGCAGCGT ATTACAGTGA CAGTTGACAG CGACAGCTAT CAGTTGCTCA AGGCATATAT	
1501	GATGTCAATA TCTCCGGTCT GGTAAAGCACA ACCATGCAGA ATGAAGCCCG TCGTCTGCGT	
1561	GCCGAACGCT GGAAAGCGGA AAATCAGGAA GGGATGGCTG AGGTCGCCCC GTTTATTGAA	
1621	ATGAACGGCT CTTTTGCTGA CGAGAACAGG GACTGGTGAA ATGCAGTTTA AGGTTTACAC	
1681	CTATAAAAGA GAGAGCCGTT ATCGTCTGTT TGTGGATGTA CAGAGTGATA TTATTGACAC	
1741	GCCCCGGCGA CGGATGGTGA TCCCCCTGGC CAGTGCACGT CTGCTGTGAG ATAAAGTCTC	
1801	CCGTGAACCT TACCCGGTGG TGCAATATCGG GGATGAAAGC TGGCGCATGA TGACCACCGA	
1861	TATGGCCAGT GTGCCGGTCT CCGTTATCGG GGAAGAAGTG GCTGATCTCA GCCACCGCGA	
1921	AAATGACATC AAAAACGCCA TTAACCTGAT GTTCTGGGGA ATATAAATGT CAGGCTCCGT	
1981	TATACACAGC CAGTCTGCAG GTCGACCATA GTGACTGGAT ATGTTGTGTT TTACAGTATT	
2041	ATGTAGTCTG TTTTTTATGC AAAATCTAAT TTAATATATT GATATTTATA TCATTTTACG	
2101	TTTCTCGTTC AGCTTTCTTG TACAAAGTGG TTCGATATCG GTACCGAGCT CGAATTCAC	
2161	GGCCGTCGTT TTACAACGTC GTGACTGGGA AAACCCTGGC GTTACCCAAC TTAATCGCCT	
2221	TGCAGCACAT CCCCCTTTTC CCAGCTGGCG TAATAGCGAA GAGGCCCGCA CCGATCGCCC	
2281	TTCCCAACAG TTGCGCAGCC TGAATGGCGA ATGGCGCCTG ATGCGGTATT TTCTCCTTAC	
2341	GCATCTGTGC GGTATTTTAC ACCGCATATG GTGCACTCTC AGTACAATCT GCTCTGATGC	
2401	CGCATAGTTA AGCCAGCCCC GACACCCGCG AACACCCGCT GACGCGCCCT GACGGGCTTG	
2461	TCTGCTCCCG GCATCCGCTT ACAGACAAGC TGTGACCGTC TCCGGGAGCT GCATGTGTCA	
2521	GAGGTTTTCA CCGTCATCAC CGAAACGCGC GAGACGAAAG GGCCTCGTGA TACGCCTATT	
2581	TTTATAGGTT AATGTCATGA TAATAATGGT TTCTTAGACG TCAGGTGGCA CTTTTCGGGG	
2641	AAATGTGCGC GGAACCCCTA TTTGTTTATT TTTCTAAATA CATTCAAATA TGTATCCGCT	
2701	CATGAGACAA TAACCTGAT AAATGCTTCA ATAATATTGA AAAAGGAAGA GTATGAGTAT	
2761	TCAACATTTT CGTGTGCGCC TTATTCCCTT TTTTGCGGCA TTTTGCCTTC CTGTTTTTGC-	

2821 TCACCCAGAA ACGCTGGTGA AAGTAAAAGA TGCTGAAGAT CAGTTGGGTG CACGAGTGGG
2881 TTACATCGAA CTGGATCTCA ACAGCGGTAA GATCCTTGAG AGTTTTCGCC CCGAAGAACC
2941 TTTTCCAATG ATGAGCACTT TTAABGTTCT GCTATGTGGC GCGGTATTAT CCCGTATTGA
3001 CGCCGGGCAA GAGCAACTCG GTCGCCGCAT AACTATTCT CAGAATGACT TGGTTGAGTA
3061 CTCACCAGTC ACAGAAAAGC ATCTTACGGA TGGCATGACA GTAAGAGAAT TATGCAGTGC
3121 TGCCATAACC ATGAGTGATA AACTGCGGC CAACCTACTT CTGACAACGA TCGGAGGACC
3181 GAAGGAGCTA ACCGCTTTTT TGCACAACAT GGGGGATCAT GTAACGCGC TTGATCGTTG
3241 GGAACCGGAG CTGAATGAAG CCATACCAA CGACGAGCGT GACACCACGA TGCCTGTAGC
3301 AATGGCAACA ACGTTGCGCA AACTATTAACT TGGCGAACTA CTTACTCTAG CTTCGCCGCA
3361 ACAATTAATA GACTGGATGG AGGCGGATAA AGTTGCAGGA CCACTTCTGC GCTCGGCCCT
3421 TCCGGCTGGC TGGTTTATTG CTGATAAATC TGGAGCCGGT GAGCGTGGGT CTCGCGGTAT
3481 CATTGCAGCA CTGGGGCCAG ATGGTAAGCC CTCCCGTATC GTAGTTATCT ACACGACGGG
3541 GAGTCAGGCA ACTATGGATG AACGAAATAG ACAGATCGCT GAGATAGGTG CCTCACTGAT
3601 TAAGCATTGG TAACTGTCAG ACCAAGTTTA CTCATATATA CTTTAGATTG ATTTAAACT
3661 TCATTTTAA TTTAAAAGGA TCTAGGTGAA GATCCTTTT GATAATCTCA TGACCAAAT
3721 CCCTTAACGT GAGTTTTCGT TCCACTGAGC GTCAGACCCC GTAGAAAAGA TCAAAGGATC
3781 TTCTTGAGAT CCTTTTTTTC TGC CGTAAT CTGCTGCTTG CAAACAAAAA AACCACCGCT
3841 ACCAGCGGTG GTTTGTTTGC CGGATCAAGA GCTACCAACT CTTTTTCCGA AGGTAAGTGG
3901 CTTCAGCAGA GCGCAGATAC CAAATACTGT CTTCTAGTG TAGCCGTAGT TAGGCCACCA
3961 CTTCAAGAAC TCTGTAGCAC CGCTACATA CCTCGCTCTG CTAATCCTGT TACCAGTGGC
4021 TGCTGCCAGT GCGGATAAGT CGTGTCTTAC CGGGTTGGAC TCAAGACGAT AGTTACCGGA
4081 TAAGGCGCAG CGGTCGGGCT GAACGGGGGG TTCTGTGACA CAGCCCAGCT TGGAGCGAAC
4141 GACCTACACC GAACTGAGAT ACCTACAGCG TGAGCTATGA GAAAGCGCCA CGCTTCCCGA
4201 AGGGAGAAAG GCGGACAGGT ATCCGGTAAG CGGCAGGGTC GGAACAGGAG AGCGCACGAG
4261 GGAGCTTCCA GGGGGAAACG CCTGGTATCT TTATAGTCCT GTCGGGTTTC GCCACCTCTG
4321 ACTTGAGCGT CGATTTTGT GATGCTCGTC AGGGGGGCGG AGCCTATGGA AAAACGCCAG
4381 CAACGCGGCC TTTTACGGT TCCTGGCCTT TTGCTGGCCT TTTGCTCACA TGTTCTTTCC
4441 TCGGTTATCC CCTGATTCTG TGGATAACCG TATTACCGCC TTTGAGTGAG CTGATACCGC
4501 TCGCCGCAGC CGAACGACCG AGCGCAGCGA GTCAGTGAGC GAGGAAGCGG AAGA

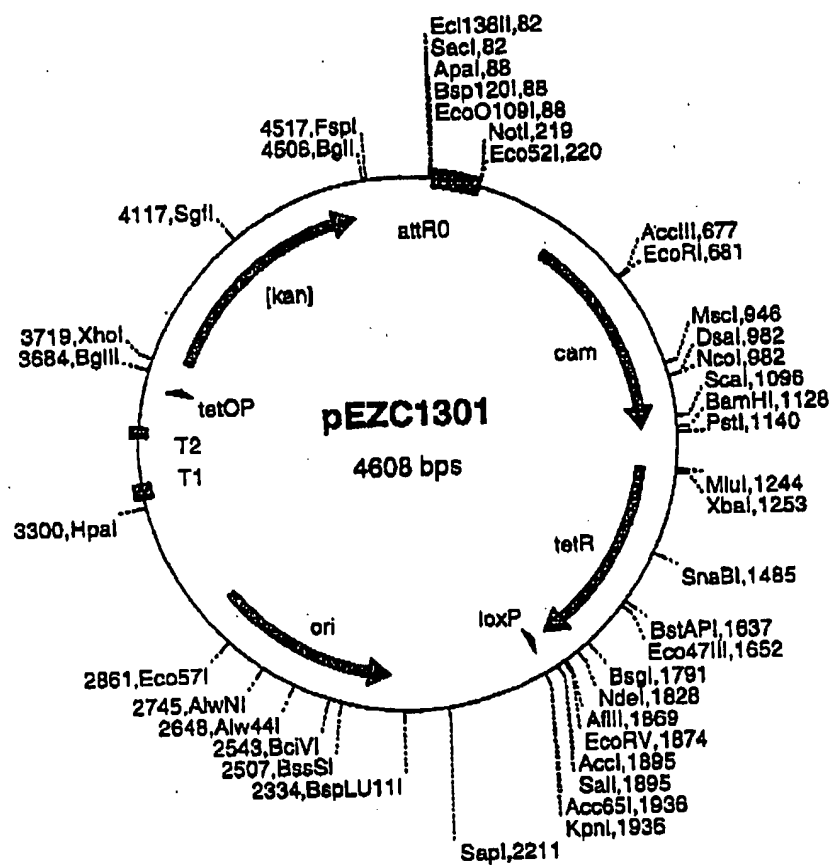


FIGURE 84

196/240

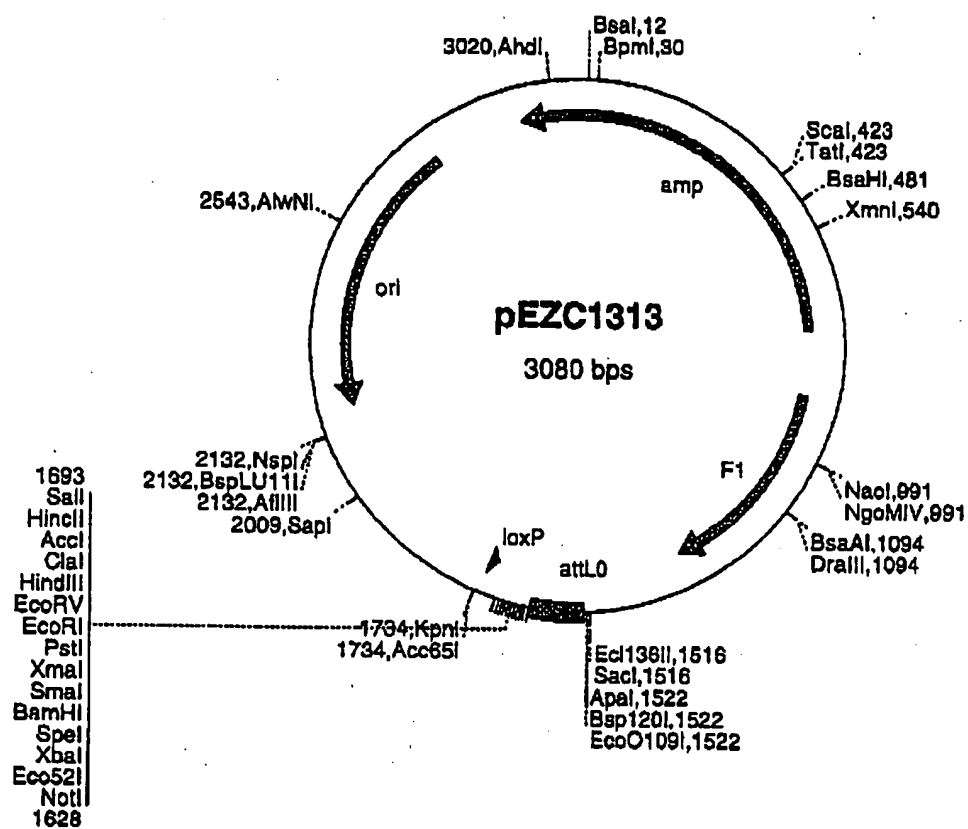


FIGURE 85

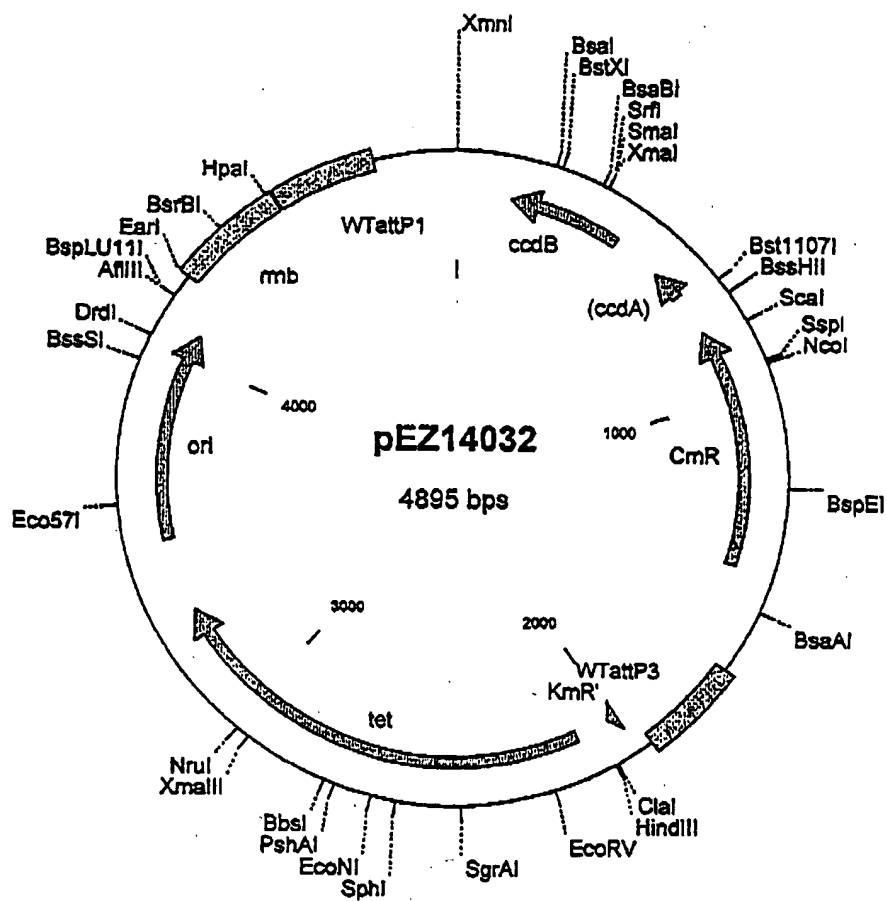


FIGURE 86

198/260

FIGURE 87

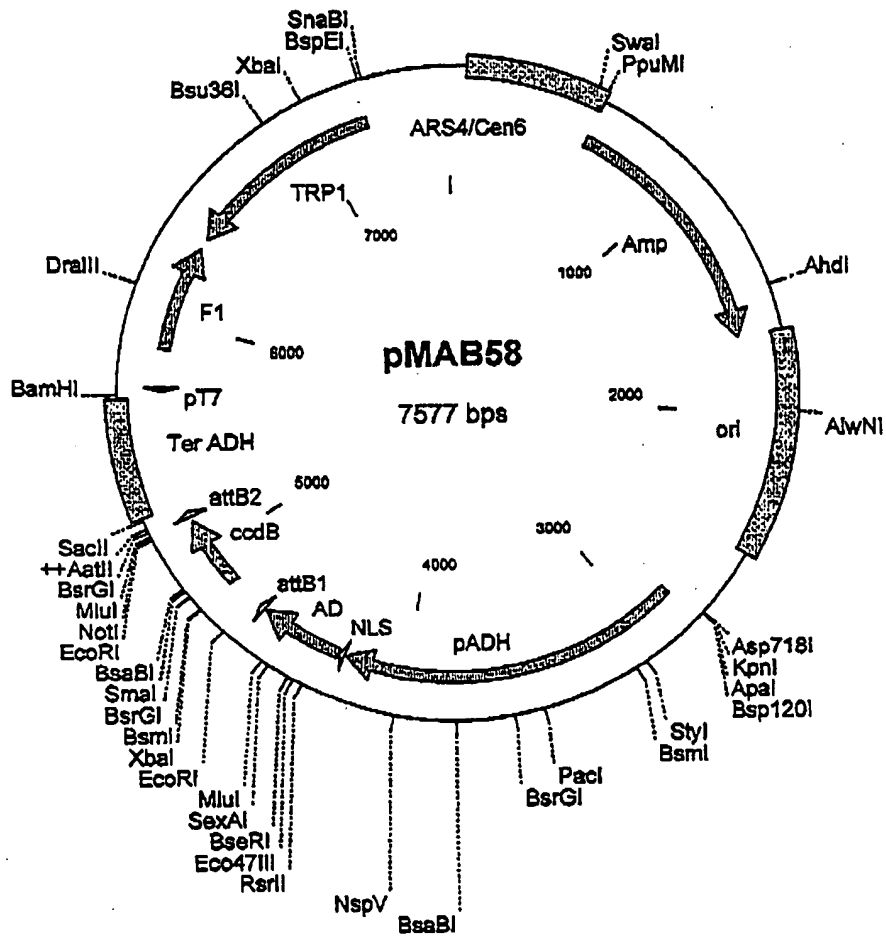
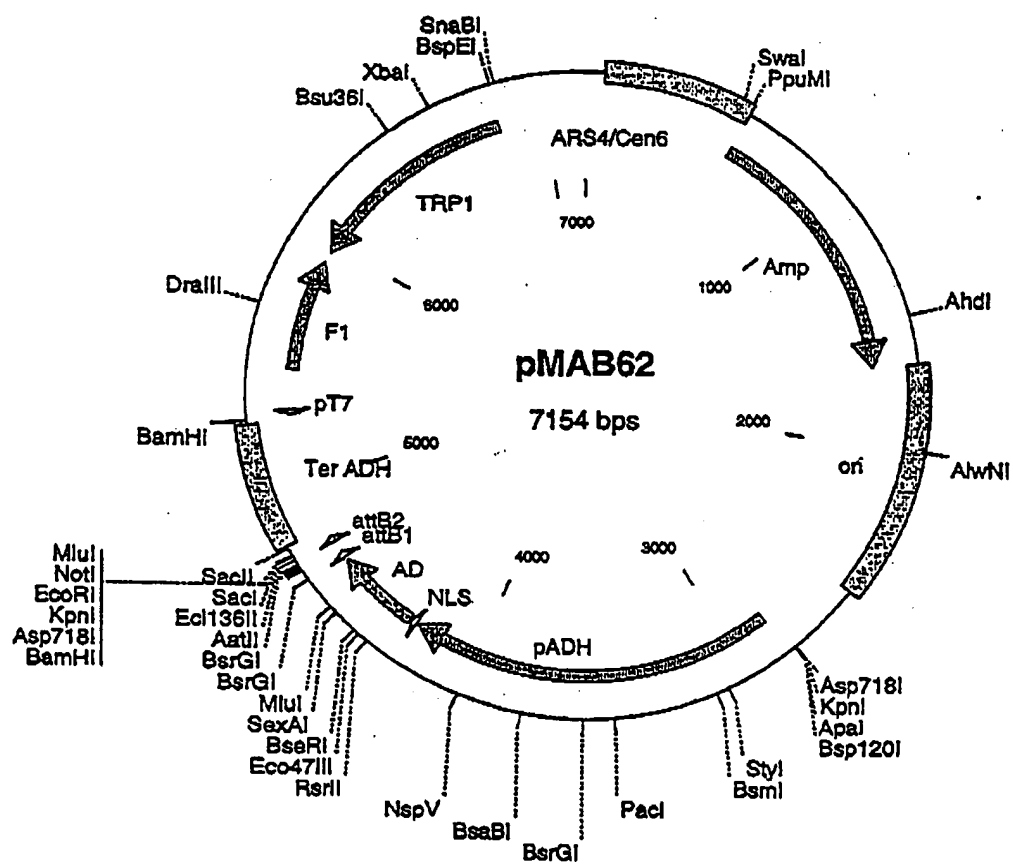
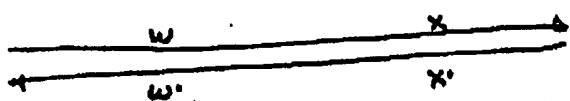


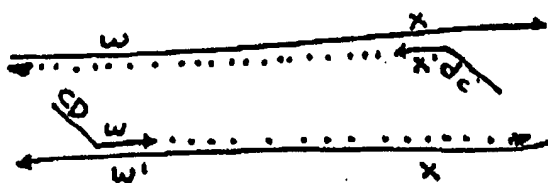
FIGURE 88



DNA to be amplified (5' → 3'):



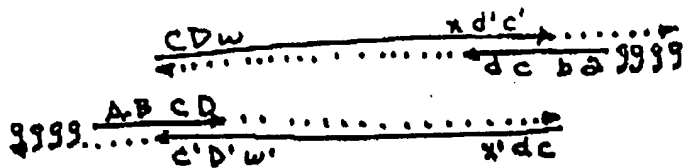
↓ Denature, anneal
hybrid primers,
↓ extend with polymerase



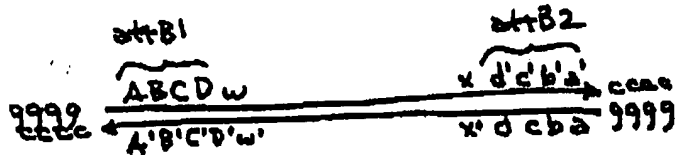
↓ amplification cycles



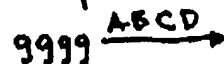
↓ Denature, anneal
attB primers,
extend with polymerase



↓ amplification cycles



attB1 primer:



attB2 primer:



Hybrid primers (part
attB, part gene
specific):



FIGURE 89

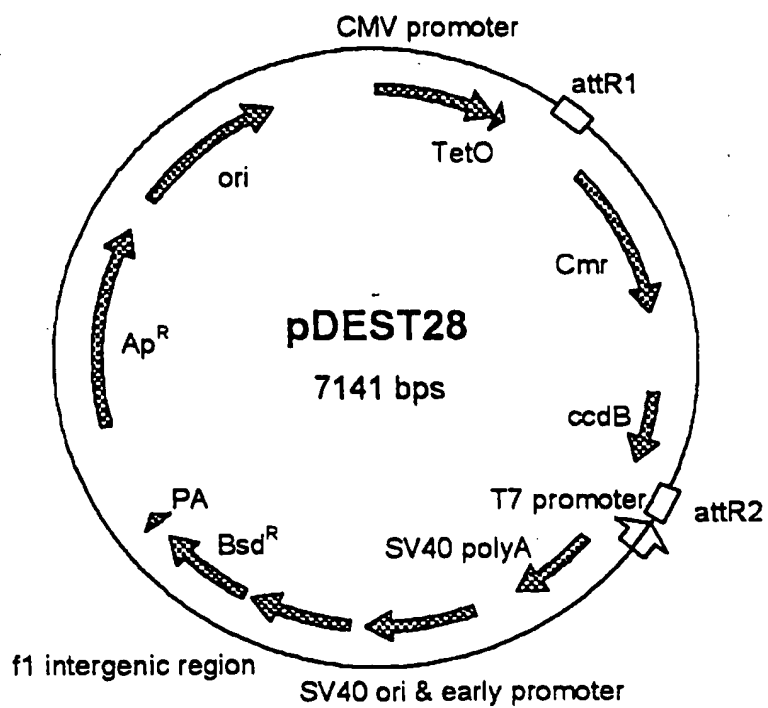


FIGURE 90A

TACGGCACCTCGACCCCAAAAACTTGATTAGGGTGATGGTTCACGTAGTGGGCCATCGC
CCTGATAGACGGTTTTTCGCCCTTTGACGTTGGAGTCCACGTTCTTTAATAGTGGACTCT
TGTTCCAACTGGAACAACACTCAACCCTATCTCGGTCTATTCTTTTGATTATAAGGGA
TTTTGCCGATTTTCGGCCTATTGGTTAAAAATGAGCTGATTAAACAAATATTTAACGCGA
ATTTTAACAAAATATTAACGTTTACAATTTTCGCCTGATGCGGTATTTTCTCCTTACGCAT
CTGTGCGGTATTTACACCGCATACGCGGATCTGCGCAGCACCATGGCCTGAAATAACCT
CTGAAAGAGGAACTTGGTTAGGTACCTTCTGAGGCGGAAAGAACCAGCTGTGGAATGTGT
GTCAGTTAGGGTGTGGAAAGTCCCCAGGCTCCCCAGCAGGCAGAAGTATGCAAAGCATGC
ATCTCAATTAGTCAGCAACCAGGTGTGGAAAGTCCCCAGGCTCCCCAGCAGGCAGAAGTA
TGCAAAGCATGCATCTCAATTAGTCAGCAACCATAGTCCCGCCCCCTAACTCCGCCCCATCC
CGCCCCCTAACTCCGCCCCAGTTCCGCCCCATTCTCGCCCCATGGCTGACTAATTTTTTTTA
TTTATGCAGAGGCCGAGGCCGCTCGGCCCTCGAGCTATTCCAGAAGTAGTGAGGAGGCT
TTTTTTGGAGGCTTAGGCTTTTGCAAAAGCTTGATTCTTCTGACACAACAGTCTCGAACT
TAAGACCATGGCCAAGCCTTTGTCTCAAGAAGAATCCACCCTCATTGAAAGAGCAACGGC
TACAATCAACAGCATCCCCATCTCTGAAGACTACAGCGTCGCCAGCGCAGCTCTCTCTAG
CGACGGCCGCATCTTCACTGGTGTCAATGTATATCATTTTACTGGGGGACCTTGTGCAGA
ACTCGTGGTGCTGGGCACTGCTGCTGCTGCGGCAGCTGGCAACCTGACTTGTATCGTCGC
GATCGGAAATGAGAACAGGGGCATCTTGAGCCCCCTGCGGACGGTGCCGACAGGTGCTTCT
CGATCTGCATCTCTGGGATCAAAGCCATAGTGAAGGACAGTGATGGACAGCCGACGGCAGT
TGGGATTCGTGAATTGCTGCCCTCTGGTTATGTGTGGGAGGGCTAAGCACTTCGTGGCCG
AGTTGCAAAATGACCGACCAAGCGACGCCCAACCTGCCATCACGATGGCCGCAATAAAATA
TCTTTATTTTCATTACATCTGTGTGTTGGTTTTTTGTGTGAATCGATAGCGATAAGGATC
CGCGTATGGTGCATCTCAGTACAATCTGCTCTGATGCCGCATAGTTAAGCCAGCCCCGA
CACCCGCCAACACCCGCTGACGCGCCCTGACGGGCTTGTCTGCTCCCGGCATCCGCTTAC
AGACAAGCTGTGACCGTCTCCGGGAGCTGCATGTGTGAGAGTTTTTCACCGTCATCACC
AAACGCGCAGACGAAAGGGCCTCGTGATACGCCATTTTATAGGTTAATGTATGATA
ATAATGGTTTTCTTAGACGTCAAGGTGGCACTTTTTCGGGGAAATGTGCGCGGAACCCCTATT
TGTTTTATTTTTCTAAATACATTCAATATGTATCCGCTCATGAGACAATAACCCCTGATAA
ATGCTTCAATAATATTGAAAAAGGAAGAGTATGAGTATCAACATTTCCGTGTGCCCTT
ATTCCCTTTTTTTCGGGCATTTTGCCTTCTGTTTTTGTCTACCCAGAAACGCTGGTGAAA
GTAAAGATGCTGAAGATCAGTTGGGTGCACGAGTGGGTACATCGAACTGGATCTCAAC
AGCGGTAAGATCCTTGAGAGTTTTTCGCCCCGAAGACGTTTTTCCAATGATGAGCACTTTT
AAAGTCTGCTATGTGGCGCGGTATTATCCCGTATTGACGCCGGGCAAGAGCACTCGGT
CGCCGCATACACTATTCTCAGAATGACTTGGTTGAGTACTACCAGTCACAGAAAAGCAT
CTTACGGATGGCATGACAGTAAGAGAATTATGCAGTGCTGCCATAACCATGAGTGATAAC
ACTGCGGCCAACTTACTTCTGACAACGATCGGAGGACCGAAGGAGCTAACCGCTTTTTTG
CACAACATGGGGGATCATGTAACCTCGCCTTGATCGTTGGGAACCGGAGCTGAATGAAGCC
ATACCAAACGACGAGCGTGACACCACGATGCCTGTAGCAATGGCAACAACGTTGCGCAA
CTATTAACCTGGCGAACTACTTACTCTAGCTTCCCGGCAACAATTAATAGACTGGATGGAG
GCGGATAAAGTTGCAGGACCACTTCTGCGCTCGGCCCTTCCGGCTGGCTGGTTATTGCT
GATAAATCTGGAGCCGGTGAGCGTGGGTCTCGCGGTATCATTGCAGCACTGGGGCCAGAT
GGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAGGCAACTATGGATGAA
CGAAATAGACAGATCGCTGAGATAGGTGCCTCACTGATTAAGCATTGGTAAGTGTGAGAC
CAAGTTTACTCATATATACTTTAGATTGATTAAAACTTCATTTTTTAATTTAAAGGATC
TAGGTGAAGATCCTTTTTGATAATCTCATGACCAAAATCCCTTAACGTGAGTTTTCGTTC
CACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTTGAGATCCTTTTTTTCTG
CGCGTAATCTGCTGCTTGCACAAAACCAACCGCTACCAGCGGTGGTTTGTGTTGCCG
GATCAAGAGCTACCAACTCTTTTTCCGAAGGTAACCTGGCTTTCAGCAGAGCGCAGATACCA
AATACTGTCTTCTAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCG
CCTACATAACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGATAAGTCCG
TGTCTTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTCCGGCTGA
ACGGGGGGTTTCGTGCACACAGCCAGCTTGGAGCGAACGACCTACACCGAACTGAGATAC
CTACAGCGTGAGCATTGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTAT
CCGTTAAGCGGCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAACGCC
TGCTATCTTTATAGTCTGTGCGGTTTCGCCACCTCTGACTTGAGCGTGCATTTTTGTGA
TGCTCGTCAGAGGGGGCGGAGCCTATGGAAAAACGCCAGCAACGCGGCCTTTTTACGGTTC
CTGGCCTTTTGTGCTGGCCTTTTGTCTCACATGTTCTTTCTGCGTTATCCCCTGATTCTGTG
GATAACCGTATTACCGCCTTTGAGTGAGCTGATACCGCTCGCCGAGCCGAACGACCGAG-

206/260

CGCAGCGAGTCAGTGAGCGAGGAAGCGGAAGAGCGCCCAATACGCAAACCGCCTCTCCCC
GCGCGTTGGCCGATTCATTAATGCAGAGCTTGCAATTCGCGCGTTTTTCAATATTATTGA
AGCATTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAAT
AAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCACCTGACGTCTAAGAAACC
ATTATTATCATGACATTAACCTATAAAAAATAGGCGTAGTACGAGGCCCTTCACTCATTA
G

FIGURE 90D

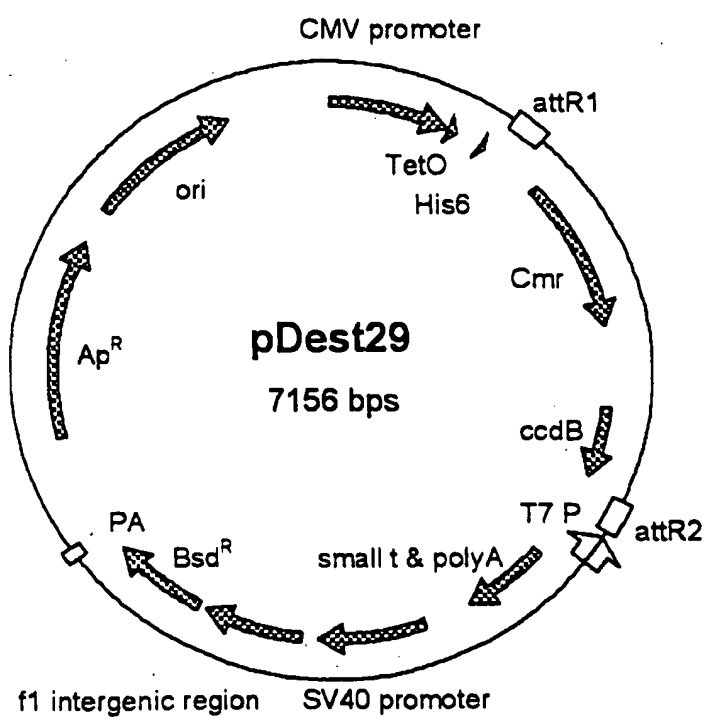


FIGURE 91 A

pDEST29 7156 bp

ATGCATGTCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCC
CGCCCATTTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCAT
TGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTAT
CATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTAT
GCCCAGTACATGACCTTATGGGACTTTCTACTTGGCAGTACATCTACGTATTAGTCATC
GCTATTACCATGGTGATGCGGTTTGGCAGTACATCAATGGGCGTGGATAGCGGTTTGAC
TCACGGGGATTTCGAAGTCTCCACCCCATTTGACGTCAATGGGAGTTTGTTTTGGCACCAA
AATCAACGGGACTTTCCAAAATGTCGTAACAACCTCGCCCCATTGACGCAAATGGGCGGT
AGGCGTGTAAGGTGGGAGGTCTATATAAGCAGAGCTCTCCCTATCAGTGATAGAGATCTC
CCTATCAGTGATAGAGATCGTCGACGAGCTCGTTTAGTGAAACCGTCAGATCGCCTGGAGA
CGCCATCCACGCTGTTTTGACCTCCATAGAAGACACCGGGACCGATCCAGCCTCCGGACC
ATGGCGTACTACCATCACCATCACCATCACCAGGTGATATCCTCGAGCCCATCACAAGT
TTGTACAAAAAGCTGAACGAGAAACGTAAATGATATAAAATATCAATATATTAAATTAG
ATTTTGCATAAAAAACAGACTACATAACTGTAAAAACACAATATCCAGTCACTATGG
CGGCCGATTAGGCACCCCGAGGCTTTACACTTTTATGCTTCCGGCTCGTATAATGTGTGA
TTTTGAGTTAGGATCCGGCGAGATTTTCAGGAGCTAAGGAAGCTAAAATGGAGAAAAAAA
TCACTGGATATACCACCGTTGATATATCCCAATGGCATCGTAAAGAACATTTTGAGGCAT
TTCAGTCAGTTGCTCAATGTACCTATAACCAGACCGTTCAGCTGGATATTACGGCCTTTT
TAAAGACCGTAAAGAAAAATAAGCACAAGTTTTATCCGGCCTTTATTACATTCTTGCCC
GCCTGATGAATGCTCATCCGGAATTCGGTATGGCAATGAAAGACGGTGAGCTGGTGATAT
GGGATAGTGTTTACCCCTTGTTACACCGTTTTCCATGAGCAAACCTGAAACGTTTTTCATCGC
TCTGGAGTGAATACCACGACGATTTCCGGCAGTTTCTACACATATATTGCAAGATGTGG
CGTGTTACGGTGAAAAACCTGGCCTATTTCCCTAAAGGTTTTATTGAGAATATGTTTTTCG
TCTCAGCCAATCCCTGGGTGAGTTTACCAGTTTTGATTTAAACGTGGCCAATATGGACA
ACTTCTTCGCCCCCGTTTTACCATGGGCAAATATTATACGCAAGGCGACAAGGTGCTGA
TGCCGCTGGCGATTACAGGTTTCATCATGCCGTCTGTGATGGCTTCCATGTCCGCAGAAATGC
TTAATGAATTACAACAGTACTGCGATGAGTGGCAGGGCGGGCGTAAACCGGTGGATCCG
GCTTACTAAAAGCCAGATAACAGTATGCGTATTTGCGCGCTGATTTTTGCGGTATAAGAA
TATATAGTATGTATACCCGAAGTATGTCAAAAAGAGGTGTGCTATGAAGCAGCGTAT
TACAGTGACGTTGACAGCGACAGCTATCGTTGCTCAAGGCATATATGATGTCAATATC
TCCGGTCTGGTAAGCACAACCATGCAGAATGAAGCCCGTCTGCTGCGTGCCGAACGCTGG
AAAGCGGAAATCAGGAAGGGATGGCTGAGGTGCGCCCGGTTTTATTGAAATGAACGGCTCT
TTTGCTGACGAGAACAGGGACTGGTGAAATGCAGTTTAAAGGTTTACACCTATAAAAGAGA
GAGCCGTTATCGTCTGTTTGTGGATGTACAGAGTGATATTATTGACACGCCCCGGGCGACG
GATGGTGATCCCCCTGGCCAGTGCACGTCTGCTGTGATATAAGTCTCCCGTGAACCTTTA
CCCCGTGGTGATATCGGGGATGAAAGCTGGCGCATGATGACCACCGATATGGCCAGTGT
GCCGTTCTCCGTTATCGGGGAAGAAGTGGTGTGATCTCAGCCACCGCGAAAATGACATCAA
AAACGCCATTAAACCTGATGTTCTGGGGAATATAAATGTGAGGCTCCGTTATACACAGCCA
GTCTGCAGGTGACCATAGTGAAGTGGATATGTTGTGTTTTACAGTATTATGTAGTCTGTT
TTTTATGCAAAATCTAATTTAATATATTGATATTTATATCATTTTACGTTTCTCGTTTACG
CTTTCTTGATACAAAGTGGTGATGGGCGGCCGCTCTAGAGGGCCCAAGCTTACGCGTGAT
GCGAGTGCATAGCTCTCTCCCTATAGTGAGTCGATTATAAGCTAGGCACTGGCCGTCGT
TTTACAACGTCGTGACTGGGAAAACCTGCTAGCTTGGGATCTTTGTGAAGGAACCTTACTT
CTGTGGTGTGACATAATTGGACAAACTACCTACAGAGATTTAAAGCTCTAAGGTAAATAT
AAAATTTTTAAGTGTATAATGTGTTAAACTAGCTGCATATGCTTGCTGCTTGAGAGTTTT
GCTTACTGAGTATGATTTATGAAAATATTATACACAGGAGCTAGTGATTCTAATTGTTTG
TGTATTTTAGATTACAGTCCCAAGGCTCATTTTCAAGCCCCCTCAGTCTCACAGTCTGTT
CATGATCATAATCAGCCATACCACATTTGTAGAGGTTTTACTTGCTTTAAAAAACCTCCC
ACACCTCCCCCTGAACCTGAAACATAAAATGAATGCAATTGTTGTTGTTAACTTGTATT
TGCAGCTTATAATGGTTACAAATAAAGCAATAGCATCACAATTTTCACAAATAAAGCATT
TTTTTCACTGTAATCTAGTTGTGGTTTTGTCCAAACTCATCAATGTATCTTATCATGTCTG
GATCGATCCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGCT
GGCGTAATAGCGAAGAGGCCCCGACCGATCGCCCTTCCCAACAGTTGCGCAGCCTGAATG
GCGAATGGGACGCGCCCTGTAGCGGCGCATTAAAGCGCGCGGGTGTGGTGGTTACGCGCA
GCGTGACCGCTACACTTGCCAGCGCCCTAGCGCCCCGCTCCTTTTCGCTTTCTTCCCTTCT
TTCTCGCCACGTTTCGCCGGCTTTCCCGCTCAAGCTCTAAATCGGGGGCTCCCTTTAGGGT-

TCCGATTTAGTGCTTTACGGCACCTCGACCCCAAAAACTTGATTAGGGTGATGGTTTAC
GTAGTGGGCCATCGCCCTGATAGACGGTTTTTCGCCCTTTGACGTTGGAGTCCACGTTCT
TTAATAGTGGACTCTTGTTCCAAACTGGAACAACACTCAACCCTATCTCGGTCTATTCTT
TTGATTATAAGGGATTTTGCCGATTTTCGGCCTATTGGTTAAAAAATGAGCTGATTTAAC
AAATATTTAACGCGAATTTTAAACAAAATATTAACGTTTACAATTTTCGCCTGATGCGGTAT
TTTCTCCTTACGCATCTGTGCGGTATTTACACCCGCATACGCGGATCTGCGCAGCACCAT
GGCCTGAAATAACCTCTGAAAGAGGAACCTTGGTTAGGTACCTTCTGAGGCGGAAAGAACC
AGCTGTGGAATGTGTGTCACTTAGGGTGTGGAAAGTCCCCAGGCTCCCCAGCAGGCAGAA
GTATGCAAAGCATGCATCTCAATTAGTCAGCAACCAGGTGTGGAAAGTCCCCAGGCTCCC
CAGCAGGCAGAAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCATAGTCCCCGCCCC
TAACTCCGCCCCATCCGCCCCCTAACTCCGCCAGTTCGCCCATTTCTCCGCCCCATGGCT
GACTAATTTTTTTTTTATTATGTCAGAGGCCGAGGCCGCTCGGCCTCTGAGCTATTCCAGA
AGTAGTGAGGAGGCTTTTTTGGAGGCCCTAGGCTTTTGCAAAAAGCTTGATTCTTCTGACA
CAACAGTCTCGAACTTAAGACCATGGCCAAGCCTTTGTCTCAAGAAGAATCCACCCTCAT
TGAAAGAGCAACGGCTACAATCAACAGCATCCCCATCTCTGAAGACTACAGCGTCGCCAG
CGCAGCTCTCTCTAGCGACGGCCGCATCTTCACTGGTGTCAATGTATATCATTTTACTGG
GGGACCTTGTGCGAACTCGTGGTGTGGGCACTGCTGCTGCTGCGGCAGCTGGCAACCT
GACTTGATCGTTCGCGATCGGAAATGAGAAACAGGGGCATCTTGAGCCCCCTGCGACGGTG
CCGACAGGTGCTTCTCGATCTGCATCTCGGATCAAAGCCATAGTGAAGGACAGTGATGG
ACAGCCGACGGCAGTTGGGATTCTGTAATTGCTGCCCTCTGGTTATGTGTGGGAGGGCTA
AGCACTTCTGCGCCGAGTTGAAATGACCGACCAAGCGACGCCCAACCTGCCATCACGAT
GGCCGCAATAAAATATCTTTATTTTTCATTACATCTGTGTGTTGGTTTTTGTGTGAATCG
ATAGCGATAAGGATCCGCGTATGGTGCATCTCAGTACAATCTGCTCTGATGCCGCATAG
TTAAGCCAGCCCCGACACCCGCCAACACCCGCTGACGCGCCCTGACGGGCTTGTCTGCTC
CCGGCATCCGCTTACAGACAAGCTGTGACCGTCTCCGGGAGCTGCATGTGTGAGAGGTTT
TCACCGTCATCACCGAAACGCGCGAGACGAAAGGGCCTCGTGATACGCCTATTTTTATAG
GTTAATGTCATGATAATAATGGTTTCTTAGACGTGAGGTGGCACTTTTCGGGGAAATGTG
CGCGGAACCCCTATTTGTTTATTTTCTAAATACATTCAAATATGTATCCGCTCATGAGA
CAATAACCCTGATAAATGCTTCAATAATATTGAAAAAGGAAGAGTATGAGTATTCAACAT
TTCCGTGTGCGCCCTTATTCCTTTTTTTCGGCATTTCGCTTCTGTTTTGCTCACCCA
GAAACGCTGGTGAAAGTAAAGATGCTGAAGATCAGTTGGGTGCACGAGTGGGTTACATC
GAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTTCGCCCGAAGAACGTTTTCCA
ATGATGAGCACTTTTAAAGTTCTGCTATGTGGCGCGGTATTATCCCGTATTGACGCCGGG
CAAGAGCAACTCGGTGCGCCGCATACACTATTCTCAGAATGACTTGGTTGAGTACTACCA
GTCACAGAAAAGCATCTTACGGATGGCATGACAGTAAGAGAATTATGCAGTGCTGCCATA
ACCATGAGTGATAACACTGCGGCCAATTACTTCTGACAACGATCGGAGGACCGAAGGAG
CTAACCCTTTTTTTCACAACATGGGGGATCATGTAACCTCGCCTTGATCGTTGGGAACCG
GAGCTGAATGAAGCCATACCAAACGACGAGCGTGACACCACGATGCCTGTAGCAATGGCA
ACAACGTTGCGCAAACTATTAACCTGGCGAATCTTACTCTAGCTTCCCGGCAACAATTA
ATAGACTGGATGGAGGCGGATAAAGTTGACAGGACCACTTCTGCGCTCGGCCCTTCCGGCT
GGCTGGTTTTATTGCTGATAAATCTGGAGCCGGTGAGCGTGGGTCTCGCGGTATCATTGCA
GCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAG
GCAACTATGGATGAACGAAATAGACAGATCGCTGAGATAGGTGCCTCACTGATTAAGCAT
TGGTAACTGTGACACCAAGTTTACTCATATATACTTTAGATTGATTTAAAACTTCATTTT
TAATTTAAAAGGATCTAGGTGAAGATCCTTTTTGATAATCTCATGACCAAAATCCCTTAA
CGTGAGTTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTTGA
GATCCTTTTTTTCTGCGCGTAATCTGCTGTGCAACAAAAAAACCACCGCTACCAGCG
GTGGTTTTGTTGCGCGATCAAGAGCTACCAACTCTTTTTCCGAAGGTAACCTGGCTTCAGC
AGAGCGCAGATACCAAATACTGTCTTCTAGTGTAGCCGTAGTTAGGCCACCACTTCAAG
AACTCTGTAGCACCGCCTACATACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCC
AGTGGCGATAAGTCTGTCTTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCG
CAGCGGTCCGGCTGAACGGGGGGTTCGTGCACACAGCCCAGCTTGGAGCGAACGACCTAC
ACCGAACTGAGATACCTACAGCGTGAGCATTGAGAAAGCGCCACGCTTCCCGAAGGGAGA
AAGGCGGACAGGTATCCGGTAAGCGGCGAGGTCGGAACAGGAGAGCGCACGAGGGAGCTT
CCAGGGGGAAACGCCTGGTATCTTTATAGTCTGTGCGGTTTCGCCACCTCTGACTTGAG
CGTCGATTTTTGTGATGCTCGTCAGGGGGCGGAGCCTATGGAAAAACGCCAGCAACCGG
GCCTTTTTACGGTTCCTGGCCTTTTGCTGGCCTTTTGCTCACATGTTCTTTCTGCGTTA
TCCCTGATTCTGTGGATAACCGTATTACCGCCTTTGAGTGAGCTGATACCGCTCGCCGC-

FIGURE 91C

AGCCGAACGACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCGGAAGAGCGCCCAATACGC
AAACCGCCTCTCCCCGCGCGTTGGCCGATTCATTAATGCAGAGCTTGCAATTCGCGCGTT
TTTCAATATTATTGAAGCATTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAA
TGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCGAAAAGTGCCACCT
GACGTCTAAGAAACCATTATTATCATGACATTAACCTATAAAAATAGGCGTAGTACGAGG
CCCTTTCACTCATTAG

FIGURE 91D

209/240

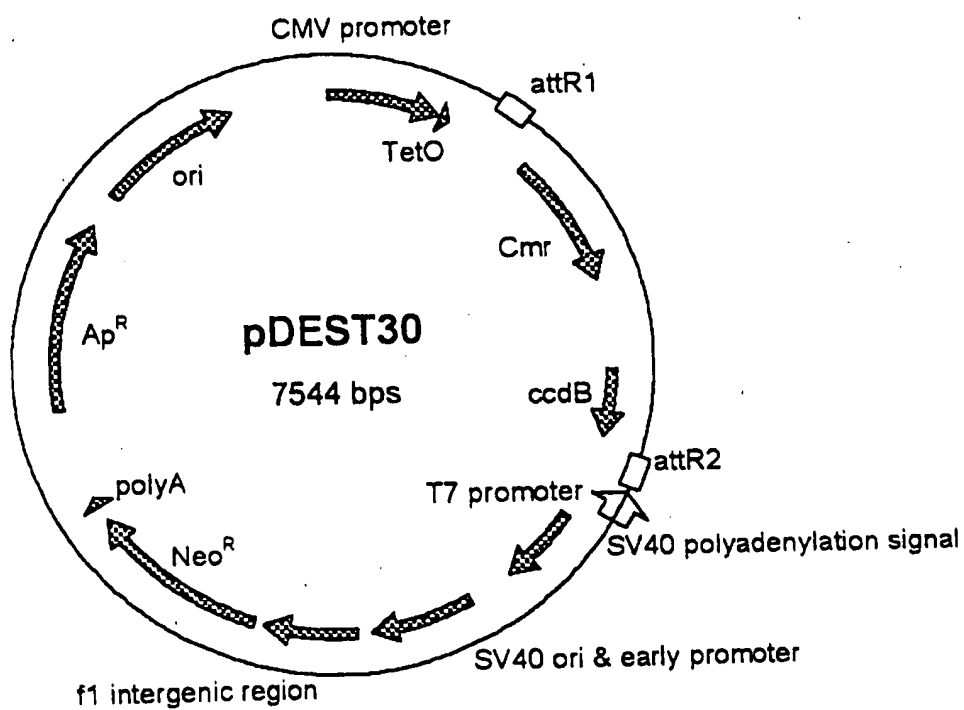


FIGURE 92A

pDEST30 7544 bp

ATGCATGTCGTTACATAACTTACGGTAAATGGCCCCGCTGGCTGACCGCCCAACGACCCC
CGCCCATTTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCAT
TGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTAT
CATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCCGCTGGCATTAT
GCCCAGTACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATC
GCTATTACCATGGTGAATGCGGTTTTGGCAGTACATCAATGGGCGTGGATAGCGGTTTTGAC
TCACGGGGATTTCCAAGTCTCCACCCCATTTGACGTCAATGGGAGTTTGTGTTTGGCACCAA
AATCAACGGGACTTTCCAAAATGTCGTAACAACCTCCGCCCCATTGACGCAAATGGGCGGT
AGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCTCCCTATCAGTGATAGAGATCTC
CCTATCAGTGATAGAGATCGTTCGACGAGCTCGTTTAGTGAACCGTCAGATCGCCTGGAGA
CGCCATCCACGCTGTTTTGACCTCCATAGAAGACACGGGACCGATCCAGCCTCCGGACT
CTAGAGGATCCCTACCGGTGATATCCTCGAGCCCATCAACAAGTTTGTACAAAAAAGCTG
AACGAGAAACGTAAATGATATAAATATCAATATATTAAATTAGATTTTGCATAAAAAAC
AGACTACATAAATACTGTAAACACAACATATCCAGTCACTATGGCGGCCGCTTAGGCAC
CCCAGGCTTTACACTTTATGCTTCCGGCTCGTATAATGTGTGGATTTTGAGTTAGGATCC
GGCGAGATTTTCAGGAGCTAAGGAAGCTAAATGGAGAAAAAATCACTGGATATACCAC
CGTTGATATATCCCAATGGCATCGTAAAGAACATTTGAGGCATTTTCACTCAGTTGCTCA
ATGTACCTATAACCAGACCGTTTCACTGGATATTACGGCCTTTTTAAAGACCGTAAAGAA
AAATAAGCACAAGTTTTATCCGGCCTTTATTACATTTCTTGCCCGCTGATGAATGCTCA
TCCGGAATTCGGTATGGCAATGAAAGACGGTGAGCTGGTATATGGGATAGTGTTCACCC
TTGTTACACCGTTTTCCATGAGCAAACTGAAACGTTTTTTCATCGCTCTGGAGTGAATACCA
CGACGATTTCCGGCAGTTTCTACACATATATTCGCAAGATGTGGCGTGTACGGTGAAAA
CCTGGCCTATTTCCCTAAAGGGTTTATTGAGAATATGTTTTTCTGCTCAGCCAATCCCTG
GGTGAGTTTACCAGTTTTGATTTAAACGTGGCCAATATGGACAACCTTCTCGCCCCCGT
TTTCACCATGGGCAATATTATACGCAAGGCGACAAGGTGCTGATGCCGCTGGCGATTCA
GGTTCATCATGCCGCTCTGTGATGGCTTCCATGTGCGCAGAATGCTTAATGAATTACAACA
GTACTGCGATGAGTGGCAGGGCGGGCGTAAAGATCTGGATCCGGCTTACTAAAAGCCAG
ATAACAGTATGCGTATTTGCGCGCTGATTTTTGCGGTATAAGAATATATACTGATATGTA
TACCCGAAGTATGTCAAAAAGAGGTGTGCTATGAAGCAGCGTATTACAGTGACAGTTGAC
AGCGACAGCTATCAGTTGCTCAAGGCATATATGATGTCAATATCTCCGGTCTGGTAAGCA
CAACCATGCAGAATGAAGCCCGTCTGCTGCGTGCCGAACGCTGGAAAGCGGAAATCAGG
AAGGGATGGCTGAGGTGCGCCGGTTTTATTGAATGAACGGCTCTTTTGCTGACGAGAACA
GGGACTGGTGAAATGCAGTTTAAGGTTTACACCTATAAAAGAGAGAGCCGTTATCGTCTG
TTTGTGGATGTACAGAGTGATATTATTGACACGCCCCGGGCGACGGATGGTGATCCCCCTG
GCCAGTGCACGTCTGCTGTGATATAAGTCTCCCGTGAACCTTTACCCGGTGGTGATATC
GGGGATGAAAGCTGGCGCATGATGACCACCGATATGGCCAGTGTGCCGGTCTCCGTTATC
GGGGAAGAAGTGGCTGATCTCAGCCACCGCGAAATGACATCAAAAACGCCATTAACCTG
ATGTTCTGGGGAATATAAATGTGAGGCTCCCTTATACACAGCCAGTCTGACAGTTCGACCA
TAGTGAAGTGGATATGTTGTGTTTTACAGTATTATGTAGTCTGTTTTTATGCAAAATCTA
ATTTAATATATTGATTTTATATCATTTTACGTTTTCTCGTTTCACTTTCTTGTACAAAGT
GGTTGATGGGCGGCCGCTCTAGAGGGCCCAAGCTTACGCGTGCATGCGACGTCATAGCTC
TCTCCCTATAGTGAGTCGTATTATAAGCTAGGCACTGGCCGTCGTTTTACAACGTCGTGA
CTGGGAAAACTGCTAGCTTGGGATCTTTGTGAAGGAACCTTACTTCTGTGGTGTGACATA
ATTGGACAAACTACCTACAGAGATTTAAAGCTCTAAGGTAAATATAAAATTTTTAAGTGT
ATAATGTGTTAACTAGCTGCATATGCTTGTGCTTGAGAGTTTTGCTTACTGAGTATGA
TTTATGAAAATATTATACACAGGAGCTAGTGATTCTAATTGTTTGTGATTTTAGATTCA
CAGTCCCAAGGCTCATTTCAAGCCCCCTCAGTCCCTCACAGTCTGTTTCATGATCATAATCAG
CCATACCACATTTGTAGAGGTTTTACTTGCTTTAAAAAACCTCCACACCTCCCCCTGAA
CCTGAAACATATAAATGAATGCAATTGTTGTTGTTAACTGTTTTATTGCAGCTTATAATGG
TTACAAATAAAGCAATAGCATCACAAATTTACAAATAAAGCATTTTTTTTCACTGCATTC
TAGTTGTGGTTTTGTCCAACTCATCAATGTATCTTATCATGTCTGGATCGATCCTGCATT
AATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGGCTATTGGCTGGCGTAATAGCGAAG
AGGCCCCGACCGATCGCCCTTCCCAACAGTTGCGCAGCCTGAATGGCGAATGGGACGCGC
CCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGTACGCGCAGCGTGACCGCTACAC
TTGCCAGCGCCCTAGCGCCCGCTCCTTTCTGCTTTCTCCCTTCTTCTCGCCACGTTCCG
CCGGCTTTCCCCGTCAAGCTCTAAATCGGGGGCTCCCTTTAGGGTTCCGATTTAGTGCTT-

FIGURE 92B

TACGGCACCTCGACCCCAAAAACTTGATTAGGGTGATGGTTCACGTAGTGGGCCATCGC
CCTGATAGACGGTTTTTCGCCCTTGACGTTGGAGTCCACGTTCTTTAATAGTGGACTCT
TGTTCCAAACTGGAACAACACTCAACCCTATCTCGGTCTATTCTTTGATTTATAAGGGA
TTTTGCCGATTTCCGCCCTATTGGTTAAAAAATGAGCTGATTTAACAAATATTTAACGCGA
ATTTTAACAAATATTAACGTTTACAATTTCCGCTGATGCGGTATTTTCTCCTTACGCAT
CTGTGCGGTATTTACACCCGCATACGCGGATCTGCGCAGCACCATGGCCTGAAATAACCT
CTGAAAGAGGAACTTGGTTAGGTACCTTCTGAGGCGGAAAGAACCAGCTGTGGAATGTGT
GTCAGTTAGGGTGTGGAAAGTCCCCAGGCTCCCCAGCAGGCAGAAGTATGCAAAGCATGC
ATCTCAATTAGTCAGCAACCAGGTGTGGAAAGTCCCCAGGCTCCCCAGCAGGCAGAAGTA
TGCAAAGCATGCATCTCAATTAGTCAGCAACCATAGTCCCGCCCCCTAACTCCGCCCCATCC
CGCCCCCTAACTCCGCCCCAGTTCGCCCCATTCTCCGCCCCATGGCTGACTAATTTTTTTTA
TTTATGCAGAGGCCGAGGCCGCTCGGCCCTCGAGCTATTCCAGAAGTAGTGAGGAGGCT
TTTTTGGAGGCCCTAGGCTTTTGCAAAAAGCTTGATTCTTCTGACACAACAGTCTCGAACT
TAAGGCTAGAGCCACCATGATTGAACAAGATGGATTGCACGCAGGTTCTCCGGCCGCTTG
GGTGGAGAGGCTATTCCGCTATGACTGGGCACAACAGACAATCGGCTGCTCTGATGCCGC
CGTGTTCCGGCTGTGACGCGCAGGGGCGCCCGGTTCTTTTTGTCAAGACCGACCTGTCCGG
TGCCCTGAATGAACTGCAGGACGAGGCGCGGCTATCGTGGCTGGCCACGACGGGCGT
TCCTTGCGCAGCTGTGCTCGACGTTGTCACTGAAGCGGGAAGGGACTGGCTGCTATTGGG
CGAAGTGCCGGGGCAGGATCTCCTGTGTCATCTCACCTTGCTCCTGCCGAGAAAGTATCCAT
CATGGCTGATGCAATGCGGCGGCTGCATACGCTTGATCCGGCTACCTGCCCATTCGACCA
CCAAGCGAAACATCGCATCGAGCGAGCACGTAACGATGGAAGCCGGTCTTGTGATCA
GGATGATCTGGACGAAGAGCATCAGGGGCTCGCGCCAGCCGAACGTTTCGCCAGGCTCAA
GGCGCGCATGCCCGACGGCGAGGATCTCGTCTGACCCATGGCGATGCCTGCTTGCCGAA
TATCATGGTGGAAAATGGCCGCTTTTCTGGATTGATCGACTGTGGCCGGCTGGGTGTGGC
GGACCGCTATCAGGACATAGCGTTGGCTACCCGTGATATTGCTGAAGAGCTTGGCGGCGA
ATGGCTGACCGCTTCTCGTGCTTTACGGTATCGCCGCTCCCGATTTCGACGCGCATCGC
CTTCTATCGCCTTCTTGACGAGTTCTTCTGAGCGGGACTCTGGGGTTCGAAATGACCGAC
CAAGCGACGCCCCAACCTGCCATCACGATGGCCGCAATAAAATATCTTTATTTTATTACA
TCTGTGTGTTGGTTTTTTGTGTGAATCGATAGCGATAAGGATCCGCGTATGGTGCATCT
CAGTACAATCTGCTCTGATGCCGCATAGTTAAGCCAGCCCCGACACCCGCCAACACCCGC
TGACGCGCCCTGACGGGCTTGTCTGCTCCCGGCATCCGCTTACAGACAAGCTGTGACCGT
CTCCGGGAGCTGCATGTGTGAGAGGTTTACCCTCATCACCGAAACGCGCGAGACGAAA
GGGCTCGTGATACGCCCTATTTTATAGGTTAATGTGATGATAATAATGGTTTTCTTAGAC
GTCAGGTGGCACTTTTCCGGGAAATGTGCGCGGAACCCCTATTTGTTTTATTTTCTAAAT
ACATTCAAATATGTATCCGCTCATGAGACAATAACCCTGATAAATGCTTCAATAATATTG
AAAAAGGAAGAGTATGAGTATTCAACATTTCCGTGTGCGCCTTATTCCTTTTTTTCGGGC
ATTTTGCCTTCTGTTTTTTGCTCACCCAGAAACGCTGGTGAAAGTAAAGATGCTGAAGA
TCAGTTGGGTGCACGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGA
GAGTTTTTCGCCCCGAAGAACGTTTTCCAATGATGAGCACTTTTAAAGTTCTGCTATGGG
CGCGGTATTATCCCGTATTGACGCCGGCAAGAGCAACTCGGTGCGGCATACACTATTC
TCAGAATGACTTGGTTGAGTACTACCAGTCAAGAAAAGCATCTTACGGATGGCATGAC
AGTAAGAGAATTATGCAGTGCTGCCATAACCATGAGTGATAACACTGCGGCCAACTTACT
TCTGACAACGATCGGAGGACCGAAGGAGCTAACCGCTTTTTTGCACAACATGGGGGATCA
TGTAACCTCGCCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAACGACGAGCG
TGACACCACGATGCCTGTAGCAATGGCAACAACGTTGCGCAAACTATTAAGTGGCGAAGT
ACTTACTCTAGCTTCCCGGCAACAATTAATAGACTGGATGGAGGCGGATAAAGTTGACAG
ACCACTTCTGCGCTCGGCCCTTCCGGCTGGCTGGTTTTATTGCTGATAAATCTGGAGCCGG
TGAGCGTGGGTCTCGCGGTATCATTTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTAT
CGTAGTTATCTACACGACGGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGATCGC
TGAGATAGGTGCCTCACTGATTAAGCATTGGTAAGTGTGACACCAAGTTTACTCATATAT
ACTTTAGATTGATTTAAAACTTCATTTTTTAATTTAAAAGGATCTAGGTGAAGATCCTTTT
TGATAATCTCATGACCAAAATCCCTTAACGTGAGTTTTTCGTTCCACTGAGCGTCAGACCC
CGTAGAAAAGATCAAAGGATCTTCTTGAGATCCTTTTTTCTGCGCGTAATCTGCTGCTT
GCAAACAAAAAACACCGCTACCAGCGGTGGTTTTGTTTGCCGGATCAAGAGCTACCAAC
TCTTTTTCCGAAGGTAAGTGGCTTACGAGAGCGCAGATACCAATACTGTCTTCTAGT
GTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCT
GCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGATAAGTCTGTCTTACCGGGTTGGA
CTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTGGGCTGAACGGGGGGTTCGTGCAC-

ACAGCCCAGCTTGGAGCGAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCATTG
AGAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGCGGCAGGGT
CGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAACGCCTGGTATCTTTATAGTCC
TGTCGGGTTTTCGCCACCTCTGACTTGAGCGTCGATTTTTGTGATGCTCGTCAGGGGGGCG
GAGCCTATGGAAAAACGCCAGCAACGCGGCCTTTTTACGGTTCCTGGCCTTTTGCTGGCC
TTTTGCTCACATGTTCTTTCCTGCGTTATCCCCTGATTCTGTGGATAACCGTATTACCGC
CTTTGAGTGAGCTGATACCGCTCGCCGCAGCCGAACGACCGAGCGCAGCGAGTCAGTGAG
CGAGGAAGCGGAAGAGCGCCCAATACGCAAACCGCCTCTCCCCGCGCGTTGGCCGATTCA
TTAATGCAGAGCTTGCAATTCGCGCGTTTTTCAATATTATTGAAGCATTATCAGGGTTA
TTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTC
GCGCACATTTCCCCGAAAAGTGCCACCTGACGTCTAAGAAACCATTATTATCATGACATT
AACCTATAAAAATAGGCGTAGTACGAGGCCCTTCACTCATTAG

FIGURE 92D

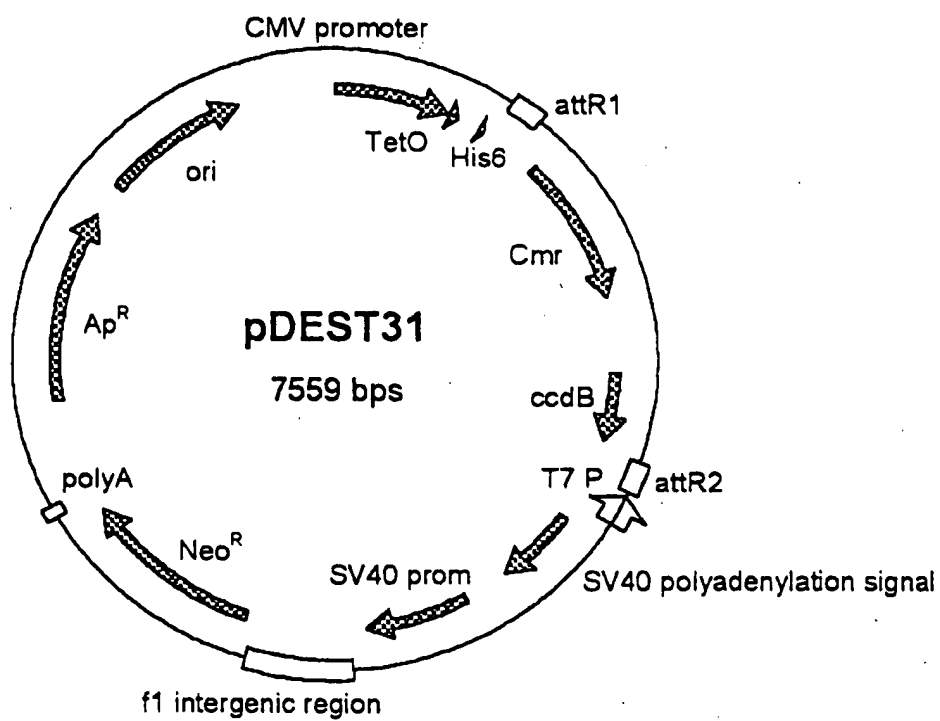


FIGURE 93A

214/240
pDEST31 7559 bp

ATGCATGTCGTTACATAACTTACGGTAAATGGCCCGCTGGCTGACCGCCCAACGACCCC
CGCCCATTTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCAT
TGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTAT
CATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCTGGCATTAT
GCCCAGTACATGACCTTATGGGACTTTTCCTACTTGGCAGTACATCTACGTATTAGTCATC
GCTATTACCATGGTGATGCGGTTTTGGCAGTACATCAATGGGCGTGGATAGCGGTTTGAC
TCACGGGGATTTCCAAGTCTCCACCCATTGACGTCAATGGGAGTTTTGTTTTGGCACCAA
AATCAACGGGACTTTCCAAAATGTCGTAACAACCTCCGCCCCATTGACGCAAATGGGCGGT
AGGCGTGACGGTGGGAGGTCTATATAAGCAGAGCTCTCCCTATCAGTGATAGAGATCTC
CCTATCAGTGATAGAGATCGTCGACGAGCTCGTTTAGTGAACCGTCAGATCGCCTGGAGA
CGCCATCCACGCTGTTTTGACCTCCATAGAAGACACCGGGACCGATCCAGCCTCCGGACC
ATGGCGTACTACCATCACCATCACCATCACACCGGTGATATCCTCGAGCCCATCACAAGT
TTGTACAAAAAAGCTGAACGAGAAACGTAATAATGATATAAATATCAATATATTAAATTAG
ATTTTGCATAAAAAACAGACTACATAACTGTAAACACAACATATCCAGTCACTATGG
CGGCCGATTAGGCACCCAGGCTTTACACTTTATGCTTCCGGCTCGTATAATGTGTGGA
TTTTGAGTTAGGATCCGGCGAGATTTTCAGGAGCTAAGGAAGCTAAAAATGGAGAAAAAA
TCACTGGATATACCACCGTTGATATATCCCAATGGCATCGTAAAGAACATTTTGAGGCAT
TTCAGTCAGTTGCTCAATGTACCTATAACACGACCGTTTCAGCTGGATATTACGGCCTTTT
TAAAGACCGTAAAGAAAAATAAGCAAGTTTTATCCGGCCTTTATTACATTCTTGCCC
GCCTGATGAATGCTCATCCGGAATTCGGTATGGCAATGAAAGACGGTGAGCTGGTGATAT
GGGATAGTGTTACCCCTTGTTACACCGTTTTCCATGAGCAAACCTGAAACGTTTTTCATCGC
TCTGGAGTGAATACCACGACGATTTCCGGCAGTTTTCTACACATATATTCGCAAGATGTGG
CGTGTTACGGTGAAACCTGGCCTATTTCCCTAAAGGGTTTATTGAGAATATGTTTTCG
TCTCAGCCAATCCCTGGGTGAGTTTCACCAAGTTTGTATTAAACGTGGCCAATATGGACA
ACTTCTTCGCCCCCGTTTTCCCATGGGCAATATATTATACGCAAGGCGACAAGGTGCTGA
TGCCGCTGGCGATTACAGTTTCATCATGCGCTGTGTATGGCTTCCATGTGCGCAGAATGC
TTAATGAATTACAACAGTACTGCGATGAGTGGCAGGGCGGGCGTAAACCGGTGGATCCG
GCTTACTAAAAGCCAGATAACAGTATGCGTATTTGCGCGCTGATTTTTTGCGGTATAAGAA
TATATACTGATATGTATACCCGAAGTATGTCAAAAAGAGGTGTGCTATGAAGCAGCGTAT
TACAGTGACAGTTGACAGCGACAGCTATCAGTTGCTCAAGGCATATATGATGTCAATATC
TCCGGTCTGGTAAGCACAACCATGCAGAATGAAGCCCGTCGTCTGCGTGCCGAACGCTGG
AAAGCGGAAAATCAGGAAGGGATGGCTGAGGTGCGCCGGTTTTATTGAAATGAACGGCTCT
TTTGCTGACGAGAACAGGGACTGGTGAAATGACAGTTTAAAGGTTTACACCTATAAAAGAGA
GAGCCGTTATCGTCTGTTTTGTGGATGTACAGAGTGATATTATTGACACGCGCGGGCGACG
GATGGTGATCCCCCTGGCCAGTGACGCTGCTGTGATGATAAAGTCTCCCGTGAACCTTA
CCCCGTGGTGATATCGGGGATGAAAGCTGGCGCATGATGACCACCGATATGGCCAGTGT
GCCGGTCTCCGTTATCGGGGAAGAGTGGCTGATCTCAGCCACCGCGAAAATGACATCAA
AAACGCCATTAACTGATGTTCTGGGGAATATAAATGTGAGGCTCCGTTATACACAGCCA
GTCTGCAGGTCGACCATAGTGACTGGATATGTTGTGTTTTACAGTATTATGTAGTCTGTT
TTTTATGCAAAATCTAATTTAATATATTGATATTTATATCATTTTACGTTTCTCGTTTACG
CTTTCTTGTAACAAAGTGGTGATGGGCGGCGCTCTAGAGGGCCCAAGCTTACGCGTGCAT
GCGACGTATAGCTCTCTCCCTATAGTGAGTCTGATTATAAGCTAGGCACTGGCCGTCGT
TTTACAACGTGCTGACTGGGAAAACCTGCTAGCTTGGGATCTTTGTGAAGGAACCTTACTT
CTGTGGTGTGACATAATTGGACAACTACCTACAGAGATTTAAAGCTCTAAGGTAAATAT
AAAATTTTTAAGTGTATAATGTGTTAAACTAGCTGCATATGCTTGCTGCTTGAGAGTTTT
GCTTACTGAGTATGATTTATGAAAATATTATACACAGGAGCTAGTGATTCTAATTGTTTG
TGATTTTATAGATTCACAGTCCCAAGGCTCATTTTCAGGCCCCCTCAGTCCTCACAGTCTGTT
CATGATCATAATCAGCCATACACATTTGTAGAGGTTTTACTTGCTTTAAAAAACCTCCC
ACACCTCCCCCTGAACCTGAAACATAAAATGAATGCAATTGTTGTTGTTAACTTGTTTAT
TGCAGCTTATAATGGTTACAAATAAAGCAATAGCATCACAATTTTCAAAATAAAGCATT
TTTTTCACTGCATTCTAGTTGTGGTTTTGTCCAAACTCATCAATGTATCTTATCATGTCTG
GATCGATCCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTTGCGTATTGGCT
GGCGTAATAGCGAAGAGGCGCCGACCGATCGCCCTTCCCAACAGTTGCGCAGCCTGAATG
GCGAATGGGACGCGCCCTGTAGCGGCGCATTAAGCGCGGGGTGTGGTGGTTACGCGCA
GCGTGACCGCTACACTTGCCAGCGCCCTAGCGCCGCTCCTTTGCTTTCTCCCTTCCT
TTCTCGCCACGTTGCGCCGGCTTTCCCGTCAAGCTCTAAATCGGGGGCTCCCTTTAGGGT-

Figure 93B

TCCGATTTAGTGCTTTACGGCACCTCGACCCCAAAAACTTGATTAGGGTGATGGTTAC
GTAGTGGGCCATCGCCCTGATAGACGGTTTTTCGCCCTTTGACGTTGGAGTCCACGTTCT
TTAATAGTGGACTCTTGTTCCAACTGGAACAACACTCAACCCTATCTCGGTCTATTCTT
TTGATTTATAAGGGATTTTGCCGATTTTCGGCCTATTGGTTAAAAAATGAGCTGATTTAAC
AAATATTTAACGCGAATTTTAACAAAATATTAACGTTTACAATTTTCGCCTGATGCGGTAT
TTTCTCCTTACGCATCTGTGCGGTATTTACACCGCATAACGCGATCTGCGCAGCACCAT
GGCCTGAAATAACCTCTGAAAGAGGAACTTGGTTAGGTACCTTCTGAGGCGGAAAGAACC
AGCTGTGGAATGTGTGTGTCAGTTAGGGTGTGGAAAGTCCCCAGGCTCCCCAGCAGGCAGAA
GTATGCAAAGCATGCATCTCAATTAGTCAGCAACCAGGTGTGGAAAGTCCCCAGGCTCCC
CAGCAGGCAGAAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCATAGTCCCCGCCC
TAACTCCGCCCATCCCCGCCCTAACTCCGCCCAGTTCGCCCCATTCTCGCCCCATGGCT
GACTAATTTTTTTTATTTATGAGAGGCCGAGGCCCTCGGCCCTCTGAGCTATTCCAGA
AGTAGTGAGGAGGCTTTTTTGGAGGCCATAGGCTTTTGCAAAAAGCTTGATTCTTCTGACA
CAACAGTCTCGAACTTAAGGCTAGAGCCACCATGATTGAACAAGATGGATTGCACGCAGG
TTCTCCGGCCGCTTGGGTGGAGAGGCTATTCCGGCTATGACTGGGCACAACAGACAATCGG
CTGCTCTGATGCCGCCGTGTTCCGGCTGTGAGCGCAGGGGCGCCCGGTTCTTTTTGTCAA
GACCGACCTGTCCGGTGCCCTGAATGAACTGCAGGACGAGGCAGCGCGGCTATCGTGGCT
GGCCACGACGGGCGTTCTTGGCGAGCTGTGCTCGACGTTGTCACTGAAGCGGGAAGGGA
CTGGCTGCTATTGGGCGAAGTGCCGGGCGAGGATCTCCTGTCTCATCTCACCTTGCTCCTGC
CGAGAAAGTATCCATCATGGCTGATGCAATGGCTGCGGCGGCTGCATACGCTTGATCCGGCTAC
CTGCCCATTCGACCACCAAGCGAAACATCGCATCGAGCGAGCACGTAAGTCCGATGGAAGC
CGGTCTTGTGATCAGGATGATCTGGACGAAGAGCATCAGGGGCTCGCGCCAGCCGAAC
GTTCCGCCAGGCTCAAGGCGCGCATGCCCGACGGCGAGGATCTCGTCTGACCCATGGCGA
TGCCTGCTTGCCGAATATCATGGTGGAAAATGCGCGCTTTTCTGGATTATCGACTGTGG
CCGGCTGGGTGTGGCGGACCGCTATCAGGACATAGCGTTGGCTACCCGTGATATTGCTGA
AGAGCTTGGCGCGAATGGGCTGACCGCTTCTCCTGCTTTACGGTATCGCCGCTCCCGA
TTCGAGCGCATCGCCTTCTATCGCCTTCTTGACGAGTTCTTCTGAGCGGGACTCTGGGG
TTCGAAATGACCGACCAAGCGACGCCCCAACCTGCCATCACGATGGCCGCAATAAAATATC
TTTATTTTTCATTACATCTGTGTGTTGGTTTTTTGTGTGAATCGATAGCGATAAGGATCCG
CGTATGGTGCATCTCAGTACAATCTGCTCTGATGCCGCATAGTTAAGCCAGCCCCGACA
CCCCGCAACACCCGCTGACGCGCCCTGACGGGCTTGTCTGCTCCCGGCATCCGCTTACAG
ACAAGCTGTGACCGTCTCCGGGAGCTGCATGTGTGAGAGTTTTACCGTCTATCACCGAA
AATGGTTTTCTAGAGCTCAGGTGGCACTTTTCGGGGAAATGTGCGCGGAACCCCTATTG
TTTTATTTTCTAAATACATTCAAATATGTATCCGCTCATGAGACAATAACCCTGATAAAT
GCTTCAATAATATTGAAAAAGGAAGAGTATGAGTATTCAACATTTCCGTGTGCGCCCTTAT
TCCCTTTTTTGCGGCATTTTGCCTTCTGTTTTTGTCTACCCAGAAACGCTGGTGAAAGT
AAAAGATGCTGAAGATCAGTTGGGTGCACGAGTGGGTTACATCGAACTGGATCTCAACAG
CGGTAAAGATCCTTGAGAGTTTTTCGCCCCGAAGAACGTTTTCCAATGATGAGCACTTTTAA
AGTTCTGCTATGTGGCGCGGTATTATCCCGTATTGACGCCGGGCAAGAGCAACTCGTTCG
CCGCATACACTATTCTCAGAATGACTTGGTTGAGTACTCACCAGTCAAGAAAAGCATCT
TACGGATGGCATGACAGTAAGAGAATTATGCAGTGCTGCCATAACCATGAGTGATAACAC
TGCGGCCCAACTTACTTCTGACAACGATCGGAGGACCGAAGGAGCTAACCGCTTTTTTGCA
CAACATGGGGGATCATGTAACCTCGCCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCAT
ACCAAACGACGAGCGTGACACCACGATGCCTGTAGCAATGGCAACAACGTTGCGCAAAC
ATTAACCTGGCGAACTACTTACTCTAGCTTCCCGGCAACAATTAATAGACTGGATGGAGGC
GGATAAAGTTGCAGGACCACTTCTGCGCTCGGCCCTTCCGGCTGGCTGGTTTTATTGCTGA
TAAATCTGGAGCCGGTGAGCGTGGGTCTCGCGGTATCATTGCAGCACTGGGGCCAGATGG
TAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAGGCAACTATGGATGAACG
AAATAGACAGATCGCTGAGATAGGTGCCTCACTGATTAAGCATTGGTAACTGTGAGACCA
AGTTTACTCATATATACTTTAGATTGATTTAAACTTCATTTTTTAATTTAAAAGGATCTA
GGTGAAGATCCTTTTTTGATAATCTCATGACCAAAATCCCTTAACGTGAGTTTTTCGTTCCA
CTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTTGAGATCCTTTTTTCTGCG
CGTAATCTGCTGCTTGCAAAACAAAAAACCCGCTACCAGCGGTGGTTGTTTGCCGGA
TCAAGAGCTACCAACTCTTTTTCCGAAGGTAACTGGCTTCAGCAGAGCGCAGATACCAAA
TACTGCTCTTCTAGTGTAGCCGTAGTTAGGCCACCCTTCAAGAACTCTGTAGCACCGCC
TACATACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGATAAGTCTGT
TCTTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTGGGCTGAAC-

Figure 93C

GGGGGGTTTCGTGCACACAGCCCAGCTTGGAGCGAACGACCTACACCGAACTGAGATACCT
ACAGCGTGAGCATTGAGAAAGCGCCACGCTTCCCCGAAGGGAGAAAGGCGGACAGGTATCC
GGTAAGCGGCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAACGCCTG
GTATCTTTATAGTCCTGTCTGGGTTTTCGCCACCTCTGACTTGAGCGTCGATTTTTGTGATG
CTCGTCAGGGGGGCGGAGCCTATGGAAAAACGCCAGCAACGCGGCCTTTTACGGTTCCT
GGCCTTTTGCTGGCCTTTTGCTCACATGTTCTTTCCTGCGTTATCCCCTGATTCTGTGGA
TAACCGTATTACCGCCTTTGAGTGAGCTGATACCGCTCGCCGCAGCCGAACGACCGAGCG
CAGCGAGTCAGTGAGCGAGGAAGCGGAAGAGCGCCCAATACGCAAACCGCCTCTCCCCGC
GCGTTGGCCGATTCATTAATGCAGAGCTTGCAATTGCGCGTTTTTCAATATTATTGAAG
CATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAA
ACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCACCTGACGTCTAAGAAACCAT
TATTATCATGACATTAACCTATAAAAAATAGGCGTAGTACGAGGCCCTTTCCTACTCATTAG

217/240

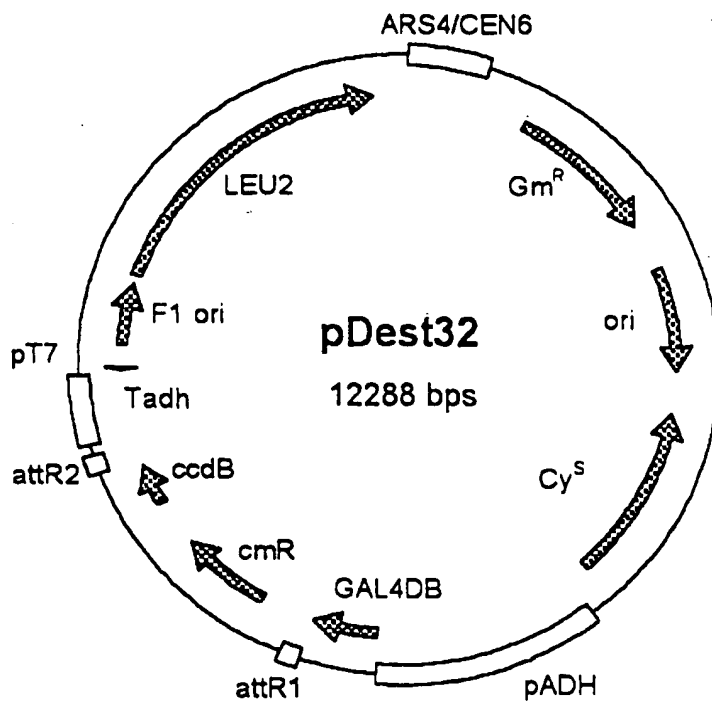


FIGURE 94A

pDEST32 12288 bp

GACGAAAGGGCCTCGTGATACGCCTATTTTTATAGGTTAATGTCATGATAATAATGGTTT
CTTAGGACGGATCGCTTGCCTGTAACCTTACACGCGCCTCGTATCTTTAATGATGGAATA
ATTTGGGAATTTACTCTGTGTTTATTTATTTTTATGTTTTGTATTGGATTTTAGAAAGT
AAATAAAGAAGGTAGAAGAGTTACGGAATGAAGAAAAAAATAAACAAAGGTTTAAAAA
ATTTCAACAAAAAGCGTACTTTACATATATATTTATTAGACAAGAAAAGCAGATTAAATA
GATATACATTTCGATTAACGATAAGTAAAATGTAAAATCACAGGATTTTCGTGTGTGGTCT
TCTACACAGACAAGATGAAACAATTCCGGCATTAAATACCTGAGAGCAGGAAGAGCAAGATA
AAAGGTAGTATTTGTTGGCGATCCCCCTAGAGTCTTTTACATCTTCGGAAAAACAAAACT
ATTTTTTCTTTAATTTCTTTTTTTTACTTTCTATTTTAAATTTATATATTTATTTAAAAA
ATTTAAATTATAATTATTTTTATAGCAGTGATGAAAAGGACCCAGGTGGCACTTTTCGG
GGAAATGTGCGCGGAACCCCTATTTGTTTATTTTTCTAAATACATTCAAATATGTATCCG
CTCATGAGACAATAACCCTGATAAATGCTTCAATAATCTGCAGTGCGCAGGGCCCCGTGTC
TCAAAATCTCTGATGTTACATTGCACAAGATAAAAATATATCATCATGAACAATAAACT
GTCTGCTTACATAAACAGTAATACAAGGGGTGTTATGAGCCATATTCAACGGGAAACGTC
TTGCTGGAGGCCCGGATTAAATTCCAACATGGATGCTGATTTATATGGGTATAAATGGGC
TCGGTAGCCAACCACTAGAACTATAGCTAGAGTCCTGGGCGAACAACGATGCTCGCCTT
CCAGAAAACCGAGGATGCGAACCCTTCTATCCGGGGTCAACACCCCGCAAGCGCCGCG
ACGGCCGAGGTCTTCCGATCTCCTGAAGCCAGGGCAGATCCGTGCACAGCACCTTGCCGT
AGAAGAACAGCAAGGCCGCAATGCCTGACGATGCGTGGAGACCGAAACCTTGCGCTCGT
TCGCCAGCCAGGACAGAAATGCCTCGACTTCGCTGCTGCCCAAGGTTGCCGGGTGACGCA
CACCGTGAAACGGATGAAGGCACGAACCCAGTTGACATAAGCCTGTTCCGTTCTGTAAAC
TGTAATGCAAGTAGCGTATGCGCTCACGCAACTGGTCCAGAACCTTGACCGAACGCGCG
GTGGTAACGGCGCAGTGGCGGTTTTTCATGGCTTGTATGACTGTTTTTTGTACAGTCTA
TGCTTCGGGCATCCAAGCAGCAAGCGCGTTACGCGGTGGGTGATGTTTGATGTTATGGA
GCAGCAACGATGTTACGCGCAGCAACGATGTTACGCGCAGGGCAGTCCGCCCTAAAAA
AAGTTAGGTGGGCTCAAGTATGGGCATCATTGCGCATGTAGGCTCGGCCCTGACCAAGTC
AAATCCATGCGGGCTGCTCTTGATCTTTTTCGGTGCTGAGTTCGGAGACGTAGCCACCTAC
TCCCAACATCAGCCGGACTCCGATTACCTCGGGAACTTGCTCCGTAGTAAGACATTCATC
GCGCTTGCTGCCTTCGACCAAGAAGCGGTTGTTGGCGCTCTCGCGGCTTACGTTCTGCC
AGGTTTGAGCAGCCCGTAGTGAGATCTATATCTATGATCTCGCAGTCTCCGGCGAGCAC
CGGAGGCAGGGCATTGCCACCGCGCTCATCAATCTCCTCAAGCATGAGGCCAACGCGCTT
GGTGCTTATGTGATCTACGTGCAAGCAGATTACGGTGACGATCCCGCAGTGGCTCTCTAT
ACAAAGTTGGGCATACGGGAAGAAGTGATGCACTTTGATATCGACCCAAGTACCGCCACC
TAACAATTGTTCAAGCCGAGATCGGCTTCCCGGCCTAATAGGTTGTATTGATGTTGGAC
GAGTCGGAATCGCAGACCGATACAGGATCTTGCCATCCTATGGAACCTGCCTCGGTGAGT
TTTCTCCTTCATTACAGAAACGGCTTTTTCAAAAATATGGTATTGATAATCCTGATATGA
ATAAATTGCAGTTTCATTTGATGCTCGATGAGTTTTTCTAATCAGAATTGTTAATTGGT
TGTAACACTGGCAGAGCATTACGCTGACTTGACGGGACGGCGNCATGACCAAAATCCCTT
AACGTGAGTTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTT
GAGATCCTTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAAAACACCGCTACCAG
CGGTGGTTTGTGTTGCCGGATCAAGAGCTACCAACTCTTTTTCCGAAGGTAACCTGGCTTCA
GCAGAGCGCAGATACCAATACTGTCCTTCTAGTGAGCCGTAGTTAGGCCACCACTTCA
AGAACTCTGTAGCACCGCTACATACCTCGCTCTGCTAATCCTGTTACAGTGGCTGCTG
CCAGTGGCGATAAGTCGTGCTTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGG
CGCAGCGGTCCGGGCTGAACGGGGGGTTTCGTGCACACAGCCAGCTTGGAGCGAACGACCT
ACACCGAACTGAGATACCTACAGCGTGAGCATTGAGAAAAGCGCCACGCTTCCCGAAGGGA
GAAAGGCGGACAGGTATCCGGTAAGCGGCAGGGTCGGAACAGGAGAGCGCACGAGGGAGC
TTCCAGGGGGGAACGCCTGGTATCTTTATAGTCTGTGCGGTTTTCGCCACCTCTGACTTG
AGCGTCGATTTTTTGTGATGCTCGTCAGGGGGGGCGAGCCTATGAAAAACGCCAGCAACG
CGGCCTTTTTTACGGTTTCTTGGCCTTTTGCTGGCCTTTTGCTCACATGTTCTTCTGCGT
TATCCCCCTGATTCTGTGGATAACCGTATTACCGCCTTTGAGTGAGCTGATACCGCTCGCC
GCAGCCGAACGACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCGGAAGAGCGCCCAATAC
GCAAACCGCCTCTCCCCGCGGTTGGCCGATTCATTAATGCAGCTGGCACGACAGGTTTC
CCGACTGGAAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTACCTCACTCATTAGG
CACCCCGAGGCTTTACACTTTATGCTTCCGGCTCCTATGTTGTGTGGAATTGTGAGCGGAT
ACAATTTTACACAGGAACAGCTATGACCATGATTACGCCAAGCTCGGAATTAACCTC-

FIGURE 94B

ACTAAAGGGAACAAAAGCTGGTACCGATCCCGAGCTTTGCAAATTAAAGCCTTCGAGCGT
CCCAAAACCTTCTCAAGCAAGGTTTTTCAAGTATAATGTTACATGCGTACACGCGTCTGTAC
AGAAAAAAGAAAAATTTGAAATATAAATAACGTTCTTAATACTAACATAACTATAAAA
AAATAAATAGGGACCTAGACTTCAGGTTGTCTAACTCCTTCCTTTTCGGTTAGAGCGGAT
GTGGGGGGAGGGCGTGAATGTAAGCGTGACATAACTAATTACATGATATCGACAAAGGAA
AAGGGGCCCTGTTTACTCACAGGCTTTTTTCAAGTAGGTAATTAAGTCGTTTCTGTCTTTT
TCCTTCTTCAACCCACCAAAGGCCATCTTGGTACTTTTTTTTTTTTTTTTTTTTTTTTTT
TT
TTTTTTTTCATAGAAATAATACAGAAGTAGATGTTGAATTAGATTAAACTGAAGATATAT
AATTTATTGGAAAATACATAGAGCTTTTTGTTGATGCGCTTAAGCGATCAATTCAACAAC
ACCACCAGCAGCTCTGATTTTTTCTTCAGCCAACTTGGAGACGAATCTAGCTTTGACGAT
AACTGGAACATTTGGAATTCTACCTTACCCTAAGATCTTACCGTAACCGGCTGCCAAAGT
GTCAATAACTGGAGCAGTTTCTTAGAAGCAGATTTCAAGTATTGGTCTCTCTGTCTTC
TGGGATCAATGTCCACAATTTGTCCAAGTTCAAGACTGGCTTCCAGAAATGAGCTTGTTG
CTTGTGGAAGTATCTCATACCAACCTTACCAGAAATAACCTGGATGGTATTTATCCATGTT
AATTCTGTGGTGTATGTTGACCACCGGCCATACCTCTACCACCGGGTGCTTTCTGTGCTT
ACCGATACGACCTTTACCGGCTGAGACGTGACCTCTGTGCTTTCTAGTCTTAGTGAATCT
GGAAGGCATTTCTGATTAGTTGGATGATTGTTCTGGGATTTAATGCAAAAATCACTTAAG
AAGGAAAATCAACGGAGAAAAGCAAACGCCATCTTAAATATACGGGATACAGATGAAAGGG
TTTGAACCTATCTGGAAAATAGCATTAACAAGCGAAAAACTGCGAGGAAAATTGTTTGC
GTCTCTGCGGGCTATTACGCGCCAGAGGAAAATAGGAAAATAACAGGGCATTAGAAAA
ATAATTTTGATTTTGGTAATGTGTGGGTCTTGGTGTACAGATGTTACATTGGTTACAGTA
CTCTTGTTTTTGCTGTGTTTTTCGATGAATCTCCAAATGGTTGTTAGCACATGGAAGAG
TCACCGATGCTAAGTTATCTCTATGTAAGCTACGTGGCGTGACTTTTGATGAAGCCGCAC
AAGAGATACAGGATTGGCAACTGCAATAGAAATCTGGGGATCCCCCTCGAGATCCGGGA
TCGAAGAAATGATGGTAAATGAAATAGGAAATCAAGGAGCATGAAGGCAAAAGACAAATA
TAAGGGTCGAACGAAAATAAAGTGAAAAGTGTGATATGATGATTTGGCTTTGCGGCG
CCGAAAAAACGAGTTTACGCAATTGCACAATCATGCTGACTCTGTGGCGGACCCGCGCTC
TTGCCGGCCCCGGCGATAACGCTGGGCGTGAGGCTGTGCCCGGCGGAGTTTTTTCGCGCTG
CATTTTCCAAGGTTTACCCTGCGCTAAGGGGCGAGATTGGAGAAGCAATAAGAATGCCGG
TTGGGGTTGCGATGATGACGACCACGACAACCTGGTGTCAATTATTAAGTTGCCGAAAGAA
CCTGAGTGCAATTTGCAACATGAGTATACTAGAAGAATGAGCCAAGACTTGCGAGACGCGA
GTTTGCCGGTGGTGGCAACAATAGAGCGACCATGACCTTGAAGGTGAGACGCGCATAACC
GCTAGAGTACTTTGAAGAGGAAACAGCAATAGGGTTGCTACCAGTATAAATAGACAGGTA
CATACAACACTGGAAATGGTTGTCTGTTTGTGAGTACGCTTTCAATTCAATTTGGGTGTGCAC
TTTATTATGTTACAATATGGAAGGGAACTTTACACTTCTCCTATGCACATATATTAATTA
AAGTCCAATGCTAGTAGAGAAGGGGGTAACACCCCTCCGCGCTCTTTTCCGATTTTTTT
CTAAACCGTGGAATATTTTCGGATATCCTTTTGTGTTTCCGGGTGTACAATATGGACTTC
CTCTTTTCTGGCAACCAAACCCATACATCGGGATTCTTATAATACCTTCGTTGGTCTCCC
TAACATGTAGGTGGCGGAGGGGAGATATAACAATAGAACAGATACCAGACAAGACATAATG
GGCTAAACAAGACTACACCAATTACACTGCCTCATTGATGGTGGTACATAACGAACATAAT
ACTGTAGCCCTAGACTTGATAGCCATCATCATATCGAAGTTTCACTACCCTTTTTCCATT
TGCCATCTATTGAAGTAATAATAGGCGCATGCAACTTCTTTTCTTTTTTTTTCTTTTCTC
TCTCCCCGTTGTTGTCTCACCATATCCGCAATGCAAAAAAATGATGGAAGACACTAA
AGGAAAAAATTAACGACAAAGACAGCACCAACAGATGTCGTTGTTCCAGAGCTGATGAGG
GGTATCTTCGAACACACGAACTTTTTCTTCTTCTTCAATTCACGCACACTACTCTTAATG
AGCAACGGTATACGGCCTTCCTTCCAGTTACTTGAATTTGAAATAAAAAAAGTTTGCCGC
TTTGCTATCAAGTATAAATAGACCTGCAATTATTAATCTTTTGTTCCTCGTCATTGTTT
TCGTTCCCTTTTCTTCTTGTCTTTTTCTGCACAATATTTCAAGCTATACCAAGCATAC
AATCAACTCCAAGCTTGAAGCAAGCCTCCTGAAAGATGAAGCTACTGTCTTCTATCGAAC
AAGCATGCGATATTTGCCGACTTAAAAAGCTCAAGTGCTCCAAAGAAAAACCGAAGTGCG
CCAAGTGCTGAAGAACAACCTGGGAGTGTCGCTACTCTCCAAAACCAAAGGTCTCCGC
TGACTAGGGCACATCTGACAGAAGTGAATCAAGGCTAGAAAGACTGGAACAGCTATTTT
TACTGATTTTTCTCGAGAAGACCTTGACATGATTTTTGAAAATGGATTCTTTACAGGATA
TAAAAGCATTTGTTAACAGGATTATTTGTACAAGATAATGTGAATAAAGATGCCGTCACAG
ATAGATTGGCTTCAGTGGAGACTGATATGCCTCTAACATTGAGACAGCATAGAATAAGTG
CGACATCATCATCGGAAGAGAGTAGTAACAAAGGTCAAAGACAGTTGACTGTATCGTCGA
GGTCAATCAAACAAGTTTGTACAAAAAGCTGAACGAGAAACGTAAATGATATAAATA-

FIGURE 94C

TCAATATATTAAATTAGATTTTGCATAAAAAACAGACTACATAATACTGTAAAAACACAAC
ATATCCAGTCACTATGGCGGCGCTAAGTTGGCAGCATCACCCGACGCACTTTGCGCCGA
ATAAATACCTGTGACGGAAGATCACTTCGCAGAATAAATAAATCCTGGTGTCCCTGTTGA
TACCGGGAAGCCCTGGGCCAACTTTTGGCGAAAATGAGACGTTGATCGGCACGTAAGAGG
TTCCAACTTTTACCATAATGAAATAAGATCACTACCGGGCGTATTTTTTTGAGTTATCGAG
ATTTTCAGGAGCTAAGGAAGCTAAAATGGAGAAAAAAATCACTGGATATACCACCGTTGA
TATATCCCAATGGCATCGTAAAGAACATTTTGAGGCATTTTCAGTCAGTTGCTCAATGTAC
CTATAACAGACACGTTTCAGCTGGATATTACGGCCTTTTAAAGACCGTAAAGAAAAATAA
GCACAAGTTTTATCCGGCCTTTTATTCACATTTTGGCCGCTGATGAATGCTCATCCGGA
ATTCCGTATGGCAATGAAAGACGGTGAGCTGGTGATATGGGATAGTGTTTACCCTTGTTA
CACCGTTTTCCATGAGCAAACCTGAAACGTTTTTCATCGCTCTGGAGTGAATACCACGACGA
TTTCCGGCAGTTTCTACACATATATTGCAAGATGTGGCGTGTTACGGTGAAAACCTGGC
CTATTTCCCTAAAGGGTTTATTGAGAATATGTTTTTCGTCTCAGCCAATCCCTGGGTGAG
TTTACCAGTTTGTATTTAAACGTGGCCAATATGGACAACTTCTTCGCCCCCGTTTTTCAC
CATGGGCAAAATATTATACGCAAGGCGACAAGGTGCTGATGCCGCTGGCGATTGAGTTTCA
TCATGCCGTCTGTGATGGCTTCCATGTCCGGCAGAATGCTTAATGAATTACAACAGTACTG
CGATGAGTGCGAGGGCGGGCGTAATCTAGAGGATCCGGCTTACTAAAAGCCAGATAACA
GTATGCGTATTTGCGCGCTGATTTTTGCGGTATAAGAATATATACTGATATGTATACCCG
AAGTATGTCAAAAAGAGGTGTGCTATGAAGCAGCGTATTACAGTGACAGTTGACAGCGAC
AGCTATCAGTTGCTCAAGGCATATATGATGTCAATATCTCCGGTCTGGTAAGCACAACCA
TGCAGAATGAAGCCCGTCGTCTGCGTGCCGAACGCTGGAAAGCGGAAAATCAGGAAGGGA
TGGCTGAGGTGCCCCGTTTATTGAAATGAACGGCTCTTTTGCTGACGAGAACAGGGACT
GGTGAAATGCAGTTTAAAGTTTTACACCTATAAAAAGAGAGAGCCGTTATCGTCTGTTGTG
GATGTACAGAGTGATATTATTGACACGCCCCGGGCGACGGATGGTGATCCCCCTGGCCAGT
GCACGTCTGCTGTGAGATAAAGTCTCCCGTGAACCTTACC CGGTGGTGATATCGGGGAT
GAAAGCTGGCGCATGATGACCACCGATATGGCCAGTGTCGGGTCTCCGTTATCGGGGAA
GAAGTGGCTGATCTCAGCCACCGCGAAAATGACATCAAAAACGCCATTAACTGATGTTT
TGGGGAATATAAATGTCAGGCTCCCTTATACACAGCCAGTCTGCAGGTGACCATAGTGA
CTGATATGTTGTGTTTTACAGTATTATGATGCTGTTTTTTATGCAAAATCTAATTTAA
TATATTGATATTATATCATTTTACGTTTCTCGTTTTCAGCTTTCTTGTAACAAAGTGGTTG
ATGGCCGCTAAGTAAGTAAGACGTCGAGCTCTAAGTAAGTAACGGCCGCCACCGCGGTGG
AGCTTTGGACTTCTTCGCCAGAGGTTTGGTCAAGTCTCCAATCAAGGTTGTGGCTTGTC
TACCTTGCCAGAAATTTACGAAAAGATGGAAAAGGGTCAAATCGTTGGTAGATACGTTGT
TGACACTTCTAAATAAGCGAATTTCTTATGATTTATGATTTTTATTATTAAATAAGTTAT
AAAAAAATAAGTGTATACAAATTTTAAAGTGACTCTTAGGTTTTAAACGAAAATCTT
GTTCTTGAGTAACTCTTTCTGTAGGTGAGTTGCTTTCTCAGGTATAGCATGAGGTGCG
TCTTATTGACCACCTCTACCGGCATGCGCGAGCAAATGCCTGCAAATCGCTCCCCATTT
CACCCAATTGTAGATATGCTAACTCCAGCAATGAGTTGATGAATCTCGGTGTGTATTTTA
TGTCCTCAGAGGACAATACCTGTTGTAATCGTCTTCCACACGGATCCCAATTCGCCCTA
TAGTGAGTCGTATTACAATTCAGTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCC
TGGCGTTACCCAACTTAATCGCCTTGACGACATCCCCCTTTCCGCCAGCTGGCGTAATAG
CGAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTTGCGCAGCCTGAATGGCGAATGGAC
GCGCCCTGTAGCGGCGCATTAAAGCGCGGCGGGTGTTGGTTACGCGCAGCGTGACCGCT
ACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTGCTTTCTTCCCTTCTTCTCGCCACG
TTCGCCGCTTTCCCGCTCAAGCTCTAAATCGGGGGCTCCCTTTAGGGTTCCGATTTAGT
GCTTTACGGCACCTCGACCCCAAAAACTTGATTAGGGTGATGGTTCACGTAGTGGGCCA
TCGCCCTGATAGACGGTTTTTCGCCCTTGACGTTGGAGTCCACGTTCTTTAATAGTGGA
CTCTTGTTCCAAACTGGAACAACACTCAACCTATCTCGGTCTATTCTTTGATTTATAA
GGGATTTTGCCGATTTCCGGCCTATTGGTTAAAAAATGAGCTGATTTAACAAAAATTTAAC
GCGAATTTTAACAAAATATTAACGTTTACAATTTCTGATGCGGTATTTTCTCCTTACGC
ATCTGTGCGGTATTTTACACCGCATATCGACCGGTGAGGAGAACTTCTAGTATATCCAC
ATACCTAATATTATTGCCTTATTAATAAATGGAATCGGAACAATTACATCAAAATCCACAT
TCTCTTCAAAATCAATTGTCCTGTACTTCTTGTTCATGTGTGTTCAAAAACGTTATATT
TATAGGATAATTATACTCTATTTCTCAACAAGTAATTGGTTGTTTGGCCGAGCGGTCTAA
GGCGCCTGATTCAAGAAATATCTTGACCGCAGTTAACTGTGGGAATACTCAGGTATCGTA
AGATGCAAGAGTTCGAATCTCTTAGCAACCATTTATTTTTTCTCAACATAACGAGAACA
CACAGGGGCGCTATCGCACAGAATCAAATTCGATGACTGGAAATTTTTTGTTAATTTTCA
AGGTGCGCTGACGCATATACCTTTTCACTGAAAAATTTGGGAGAAAAAGGAAAGGTGAG-

FIGURE 94D

AGGCCGGAACCGGCTTTTCATATAGAATAGAGAAGCGTTCATGACTAAATGCTTGCAATCA
CAATACTTGAAGTTGACAATATTATTTAAGGACCTATTGTTTTTCCAATAGGTGGTTAG
CAATCGTCTTACTTTCTAACTTTTCTTACCTTTTACATTTTACGCAATATATATATATATT
TCAAGGATATACCATTTCTAATGTCTGCCCCCTATGTCTGCCCCCTAAGAAGATCGTCTGTTT
GCCAGGTGACCACGTTGGTCAAGAAATCACAGCCGAAGCCATTAAAGGTTCTTAAAGCTAT
TTCTGATGTTTCGTTCCAATGTCAAGTTTCGATTTTCGAAAATCATTTAATTGGTGGTGCTGC
TATCGATGCTACAGGTGTCCCACTTCCAGATGAGGCGCTGGAAGCCTCCAAGAAGGTTGA
TGCCGTTTTGTTAGGTGCTGTGGGTGGTCTTAAATGGGGTACCGGTAGTGTTAGACCTGA
ACAAGGTTTACTAAAAATCCGTAAAGAACTTCAATTGTACGCCAACTTAAGACCATGTAA
CTTTGCATCCGACTCTCTTTTAGACTTATCTCCAATCAAGCCACAATTTGCTAAAGGTAC
TGACTTTCGTTGTTGTCTCAGAGAATTAGTGGGAGGTATTTACTTTGGTAAGAGAAAGGAAGA
CGATGGTGTATGGTGTCTGCTTGGGATAGTGAACAATACACCGTTCCAGAAGTGCAAAGAAT
CACAAGAATGGCCGCTTTTCATGGCCCTACAAACATGAGCCACCATTGCTTATTTGGTCTCTT
GGATAAAGCTAATGTTTTGGCCTCTTCAAGATTATGGAGAAAACTGTGGAGGAAACCAT
CAAGAACGAATTCCTACATTGAAGGTTCAACATCAATTGATTGATTCTGCCGCCATGAT
CCTAGTTAAGAACCCCAACCCACCTAAATGGTATTATAATCACCAGCAACATGTTTGGTGA
TATCATCTCCGATGAAGCCTCCGTTATCCCAGGTTCTTTGGGTTTGTGGCATCTGCGTC
CTTGGCCTCTTTGCCAGACAAGAACACCGCATTTGGTTTGTACGAACCATGCCACGGTTC
TGCTCCAGATTTGCCAAAGAATAAGGTTGACCCCTATCGCCACTATCTTGTCTGTCTGCAAT
GATGTTGAAATTGTCAATTGAACCTTGCCCTGAAGAAGGTAAGGCCATTGAAGATGCAGTTAA
AAAGGTTTTGGATGCAGGTATCAGAACTGGTGATTTAGGTGGTTCCAACAGTACCACCGA
AGTCCGTGATGCTGTGCGCCGAAGAAGTTAAGAAAATCCTTGCTTAAAAAGATTCTCTTTT
TTTATGATATTTGTACATAAACTTTATAAATGAAATTCATAATAGAAACGACACGAAATT
ACAAAATGGAATATGTTTCATAGGGTAGACGAACTATATACGCAATCTACATACATTTAT
CAAGAAGGAGAAAAAGGAGGATAGTAAAGGAATACAGGTAAGCAAATTGATACTAATGGC
TCAACGTGATAAGGAAAAAGAATTGCACTTTAACATTAATATTGACAAGGAGGAGGCGAC
CACACAAAAAGTTAGGTGTAACAGAAAATCATGAACTACGATTCCTAATTTGATATTGG
AGGATTTTTCTCTAAAAAATAAATAACAATAAAAAACACTCAATGACCTGACCAT
TTGATGGAGTTTAAGTCAATACCTTCTTGAACCATTTCCTAATGGTGAAAGTTCCCTC
AAGAATTTTACTCTGTGAGAAACGGCCTTACGACGTAGTCGATATGGTGCACTCTCAGTA
CAATCTGCTCTGATGCCGCATAGTTAAGCCAGCCCCGACACCCGCCAACACCCGCTGACG
CGCCCTGACGGGCTTGTCTGCTCCCGGCATCCGCTTACAGACAAGCTGTGACCGTCTCCG
GGAGCTGCATGTGTCTCAGAGGTTTTACCGTTCATCACCGAAACGCGCGA

FIGURE 94E

222/240

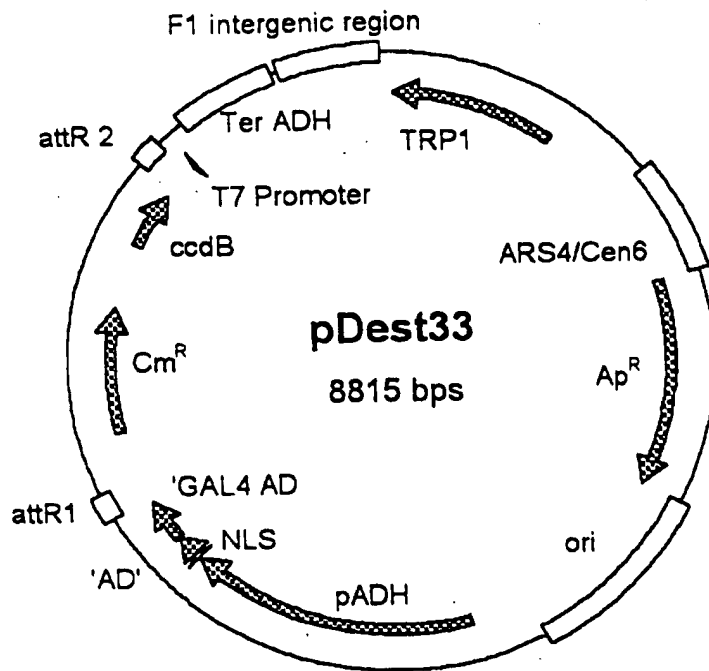


FIGURE 95A

pDEST33 8815 bp

GCCTTACGCATCTGTGCGGTATTTTCACACCGCAGGCAAGTGCACAAACAATACTTAAATA
AATACTACTCAGTAATAACCTATTTCTTAGCATTTTTGACGAAATTTGCTATTTTGTTAG
AGTCTTTTACACCATTGTCTCCACACCTCCGCTTACATCAACACCAATAACGCCATTTA
ATCTAAGCGCATCACCAACATTTTCTGGCGTCAGTCCACCAGCTAACATAAAATGTAAGC
TTTTCGGGGCTCTCTTGCCCTTCCAACCCAGTCAGAAATCGAGTTCCAATCCAAAAGTTCAC
CTGTCCCACCTGCTTCTGAATCAAACAAGGGAATAAACGAATGAGGTTTCTGTGAAGCTG
CACTGAGTAGTATGTTGCAGTCTTTTGGAAATACGAGTCTTTTAATAACTGGCAAACCGA
GGAACTCTTGGTATTCTTGCCACGACTCATCTCCATGCAGTTGGACGATATCAATGCCGT
AATCATTGACCAGAGCCAAAACATCCTCCTTAGGTTGATTACGAAACACGCCAACCAAGT
ATTTCCGGAGTGCCTGAACTATTTTATATGCTTTTACAAGACTTGAAATTTTCTTGCAA
TAACCGGGTCAATTGTTCTCTTTCTATTGGGCACACATATAATACCCAGCAAGTCAGCAT
CGGAATCTAGAGCACATTCTGCGGCCTCTGTGCTCTGCAAGCCGCAAACCTTTCACCAATG
GACCAGAACTACCTGTGAAATTAATAACAGACATACTCCAAGCTGCCTTTGTGTGCTTAA
TCACGTATACTCAGTGCTCAATAGTCACCAATGCCCTCCCTCTTGGCCCTCTCCTTTTC
TTTTTTCGACCGAATTAATTCCTAATCGGCAAAAAAGAAAAGCTCCGGATCAAGATTGT
ACGTAAGGTGACAAGCTATTTTCAATAAAGAATATCTTCCACTACTGCCATCTGGCGTC
ATAACTGCAAAGTACACATATATTACGATGCTGTCTATTAAATGCTTCCTATATTATATA
TATAGTAATGTCGTTTATGGTGCACTCTCAGTACAATCTGCTCTGATGCCGCATAGTTAA
GCCAGCCCCGACACCCGCCAACACCCGCTGACGCGCCCTGACGGGCTTGTCTGCTCCCGG
CATCCGCTTACAGACAAGCTGTGACCGTCTCCGGGAGCTGCATGTGTGAGAGGTTTTAC
CGTCATCACCGAAACGCGCGAGACGAAAGGGCCTCGTGATACGCCTATTTTTATAGGTTA
ATGTCATGATAATAATGGTTTCTTAGGACGGATCGCTTGCCCTGTAACCTACACGCGCCTC
GTATCTTTTAAATGATGGAATAATTTGGGAATTTACTCTGTGTTTATTTATTTTATGTTT
TGTATTTGGATTTTAGAAAGTAAATAAAGAGGTAGAAGAGTTACGGAATGAAGAAAAA
AAATAAACAAAGGTTTAAAAAATTTCAACAAAAAGCGTACTTTACATATATATTTATTAG
ACAAGAAAAGCAGATTAAATAGATATACATTGATTAACGATAAGTAAATGTAAATCA
CAGGATTTTTCGTGTGTTGTTCTTACACAGACAAGATGAAACAATTCGGCATTAATACCT
GAGAGCAGGAAGAGCAAGATAAAAGGTAGTATTTGTTGGCGATCCCCCTAGAGTCTTTTA
CATCTTTCGGAAAACAAAACTATTTTTCTTTAATTTCTTTTTTTACTTTCTATTTTTAA
TTTATATATTTATATTAATAAATTTAAATTATAATTATTTTTATAGCAGTGATGAAAG
GACCCAGGTGGCACTTTTCGGGGAATGTGCGCGGAACCCCTATTTGTTTATTTTTCTAA
ATACATTCAAATATGATCCGCTCATGAGACAATAACCCTGATAAATGCTTCAATAATAT
TGAAAAAGGAAGAGTATGAGTATTCAACATTTCCGTGTGCCCCCTATTCCCTTTTTTGCG
GCATTTTGCCTTCTGTTTTTGTCTACCCAGAAACGCTGGTGAAAGTAAAGATGCTGAA
GATCAGTTGGGTGCACGAGTGGGTTACATCGAAGTGGATCTCAACAGCGGTAAGATCCTT
GAGAGTTTTTCGCCCCGAAGAAGCTTTTCCAATGATGAGCACTTTTAAAGTTCTGCTATGT
GGCGCGGTATTATCCCGTATTGACGCGGGCAAGAGCAACTCGGTGCGCGCATACACTAT
TCTCAGAATGACTTGGTTGAGTACTCACCAGTCACAGAAAAGCATCTTACGGATGGCATG
ACAGTAAGAGAATTATGCAAGTGCTGCCATAACCATGAGTGATAACACTGCGGCCAACTTA
CTTCTGACAACGATCGGAGGACCGAAGGAGCTAACCGCTTTTTTTCACAACATGGGGGAT
CATGTAACCTCGCCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAACGACGAG
CGTGACACCACGATGCCTGTAGCAATGGCAACAACGTTGCGCAAACTATTAAGTGGCGAA
CTACTTACTCTAGCTTCCCGGCAACAATTAATAGACTGGATGGAGGCGGATAAAGTTGCA
GGACCACTTCTGCGCTCGGCCCTTCCGGCTGGCTGGTTTATTGCTGATAAATCTGGAGCC
GGTGAGCGTGGGTCTCGCGGTATCATTTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGT
ATCGTAGTTATCTACAGACGGGCAGTCAGGCAACTATGGATGAACGAAATAGACAGATC
GCTGAGATAGGTGCCTCACTGATTAAGCATTTGGTAACTGTCAGACCAAGTTTACTCATAT
ATACTTTAGATTGATTTAAAACTTCATTTTTTAATTTAAAGGATCTAGGTGAAGATCCTT
TTTGATAATCTCATGACCAAAATCCCTTAACGTGAGTTTTTCGTTCCACTGAGCGTCAGAC
CCCGTAGAAAAGATCAAAGGATCTTCTTGAGATCCTTTTTTTCTGCGCGTAATCTGCTGC
TTGCAACAAAAAACCACCGCTACCAGCGGTGGTTTGTGTTGCCGGATCAAGAGCTACCA
ACTCTTTTTCCGAAGGTAAGTGGCTTCAGCAGAGCGCAGATACCAAACTGTCTCTTA
GTGTAGCCGTAGTTTAGGCCACCACTTCAAGAACTCTGTAGCACCCTACATACCTCGCT
CTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTTG
GACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTGGGGCTGAACGGGGGGTTCTGTG
ACACAGCCCAGCTTGGAGCGAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCAT-

FIGURE 95B

TGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGCGGCAGG
GTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGGAACGCCTGGTATCTTTATAGT
CCTGTGGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTTGTGATGCTCGTCAGGGGGG
CCGAGCCTATGGAAAAACGCCAGCAACGCGGCCTTTTTACGGTTCCTGGCCTTTTGCTGG
CCTTTTGCTCACATGTTCTTTCTGCGTTATCCCCTGATTCTGTGGATAACCGTATTACC
GCCTTTGAGTGAGCTGATACCGCTCGCCGCGAGCCGAACGACCCGAGCGCAGCGAGTCAGTG
AGCGAGGAAGCGGAAGAGCGCCCAATACGCAAACCGCCTCTCCCCGCGCGTTGGCCGATT
CATTAAATGCAGCTGGCACGACAGGTTTCCCGACTGGAAAGCGGGCAGTGAGCGCAACGCA
ATTAATGTGAGTTACCTCACTCATTAGGCACCCCAGGCTTTACACTTTATGCTTCCGGCT
CCTATGTTGTGTGGAATTGTGAGCGGATAACAATTTACACAGGAAACAGCTATGACCAT
GATTACGCCAAGCTCGGAATTAACCCTCACTAAAGGGGAACAAAAGCTGGGTACCGGGCCC
CCCCCTCGAGATCCGGGATCGAAGAAATGATGGTAAATGAAATAGGAAATCAAGGAGCATG
AAGGCAAAAGACAAATATAAGGGTCTGAACGAAAAATAAAGTGAAAAGTGTGATATGATG
TATTTGGCTTTTGGCGCGCCGAAAAACGAGTTTACGCAATTGCACAATCATGCTGACTCT
GTGGCGGACCCGCGCTCTTGCCGGCCCGCGGATAACGCTGGGCGTGAGGCTGTGCCCGGC
GGAGTTTTTTGCGCCTGCATTTTCCAAGGTTTACCCTGCGCTAAGGGGCGAGATTGGAGA
AGCAATAAGAATGCCGTTGGGGTTGCGATGATGACGACCACGACAACCTGGTGTCATTAT
TTAAGTTGCCGAAAGAACCCTGAGTGCATTTGCAACATGAGTATACTAGAAGAATGAGCCA
AGACTTGCGAGACGCGAGTTTGCCTGGTGCGAACAATAGAGCGACCATGACCTTGAAAG
GTGAGACGCGCATAACCGCTAGAGTACTTTGAAGAGGAAACAGCAATAGGGTTGCTACCA
GTATAAATAGACAGGTACATAACAACCTGGAAATGGTTGTCTGTTTGTGAGTACGCTTTCAA
TTCATTTGGGTGTGCACTTTATTATGTTACAATATGGAAGGGAACTTTACACTTCTCCTA
TGCACATATATTAATTAAGTCCAATGCTAGTAGAGAAGGGGGGTAACACCCCTCCGCGC
TCTTTTCCGATTTTTTTCTAAACCGTGGAATATTTTCGGATATCCTTTTGTGTTTCCGGG
TGTACAATATGGACTTCTCTTTTCTGGCAACCAACCCATACATCGGGATTCTCTATAAT
ACCTTCGTTGGTCTCCCTAACATGTAGGTGGCGGAGGGGAGATATACAATAGAACAGATA
CCAGACAAGACATAATGGGCTAAACAAGACTACACCAATTACACTGCCTCATTGATGGTG
GTACATAACGAACATAACTGTAGCTAGCCCTAGACTTGATAGCCATCATCATATCGAAGTTTC
ACTACCCCTTTTCCATTTTGCCATCTATTGAAGTAATAATAGGCGCATGCAACTTCTTTTC
TTTTTTTTTCTTTTCTCTCTCCCCGTTGTTGTCTCACCATATCCGCAATGACAAAAAAA
ATGATGGAAGACACTAAAGGAAAAAATTAACGACAAAGACAGCACCAACAGATGTCGTTG
TTCCAGAGCTGATGAGGGGTATCTTCGAACACACGAACTTTTTCTTCTTCTTCAATCAG
CACACTACTCTCTAATGAGCAACGGTATACGGCCTTCTTCCAGTTACTTGAATTTGAAA
TAAAAAAGTTTGCCGCTTTGCTATCAAGTATAAATAGACCTGCAATTATTAATCTTTTG
TTTCTCGTCAATTGTTCTCGTTCCTTTTCTTCTTGTCTTTTCTGCACAATATTTCA
AGCTATACCAAGCATAACAATCAACTCCAAGCTTATGCCCAAGAAGAAGCGGAAGGTCTCG
AGCGGCGCCAATTTTAATCAAAGTGGAATATTGCTGATAGCTCATTGTCTTCACTTTC
ACTAACAGTAGCAACGGTCCGAACCTCATAACAACCTCAAACAAATTCTCAAGCGCTTTC
CAACCAATTGCCTCCTCTAACGTTTCATGATAACTTCATGAATAATGAAATCACGGCTAGT
AAAATTGATGATGGTAATAATTCAAACCACTGTCACCTGGTTGGACGGACCAAACTCGG
TATAACGCGTTTGGAACTACTACAGGGATGTTTAAATACCACTACAATGGATGATGATAT
AACTATCTATTGATGATGAAGATACCCCAACCAACCAAAAAAGAGGGTGGGTGCAAT
CAAACAAGTTTGTACAAAAAGCTGAACGAGAAACGTAAATGATATAAATATCAATATA
TTAAATTAGATTTTGCATAAAAAACAGACTACATAATACTGTAAACACAACATATCCAG
TCACTATGGCGGCCGCTAAGTTGGCAGCATCACCCGACGCACTTTGCGCCGAATAAATAC
CTGTGACGGAAGATCACTTCGAGAAATAAATAAATCCTGGTGTCCCTGTTGATACCGGGA
AGCCCTGGGCGCACTTTTGGCGAAAATGAGACGTTGATCGGCACGTAAGAGGTTCCAAC
TTCACCATAATGAAATAAGATCACTACCGGGCGTATTTTTTGTGATTATCGAGATTTTCAG
GAGCTAAGGAAGCTAAATGGAGAAAAAATCACTGGATATACCACCGTTGATATATCCC
AATGGCATCGTAAAGAACATTTTGGAGGCATTTTCAGTCAGTTGCTCAATGTACCTATAACC
AGACCGTTTCAGCTGGATATTACGGCCTTTTTTAAAGACCGTAAAGAAAAATAAGCACAAGT
TTTATCCGCGCTTTTATTCATCTTGTCCCGCCTGATGAATGCTCATCCGGAATTCGGTA
TGGCAATGAAAGACGGTGAGCTGGTGTATGGGATAGTGTTCACCCTTGTACACCGTTT
TCCATGAGCAAACTGAAACGTTTTCATCGCTCTGGAGTGAATACCACGACGATTTCCGGC
AGTTTCTACACATATATTGCAAGATGTGGCGGTGTTACGGTGAAAAACCTGGCCTATTTCC
CTAAAGGGTTTATTGAGAATATGTTTTTCTGCTCAGCCAATCCCTGGGTGAGTTTCACCA
GTTTTGATTTAAACGTGGCCAATATGGACAACCTTCTCGCCCCCGTTTTTACCATGGGCA
AATATTATACGCAAGGCGACAAGGTGCTGATGCCGCTGGCGATTACAGGTTTCATCATGCCG-

FIGURE 95C

TCTGTGATGGCTTCCATGTCCGGCAGAATGCTTAATGAATTACAACAGTACTGCGATGAGT
GGCAGGGCGGGGCGTAATCTAGAGGATCCGGCTTACTAAAAGCCAGATAACAGTATGCGT
ATTTGCGCGCTGATTTTTGCGGTATAAGAATATATACTGATATGTATACCCGAAGTATGT
CAAAAAGAGGTGTGCTATGAAGCAGCGTATTACAGTGACAGTTGACAGCGACAGCTATCA
GTTGCTCAAGGCATATATGATGTCAATATCTCCGGTCTGGTAAGCACCAACCATGCAGAAT
GAAGCCCGTCTGCTGCGTGCCGAACGCTGGAAAGCGGAAAATCAGGAAGGGATGGCTGAG
GTCGCCCCGGTTTATTGAAATGAACGGCTCTTTTGCTGACGAGAACAGGGACTGGTGAAAT
GCAGTTTAAAGTTTACACCTATAAAAGAGAGAGCCGTTATCGTCTGTTTGTGGATGTACA
GAGTGATATTATTGACACGCCCCGGGCGACGGATGGTGATCCCCCTGGCCAGTGCACGCTCT
GCTGTGAGATAAAGTCTCCCGTGAACTTTACCCGGTGGTGCATATCGGGGATGAAAGCTG
CGCATGATGACCACCGATATGGCCAGTGTGCGGGTCTCCGTTATCGGGGAAGAAGTGCC
TGTCTCAGCCACCGGAAAATGACATCAAAAACGCCATTAACTGATGTTCTGGGGAAT
ATAAATGTCAGGCTCCGTTATACACAGCCAGTCTGCAGGTGACCATAGTGAAGTGGATAT
GTTGTGTTTTACAGTATTATGTAGTCTGTTTTTTATGCAAAATCTAATTTAATATATTGA
TATTTATATCATTTTACGTTTCTCGTTTCTGTTTCTTGTACAAAGTGGTTTGTATGGCCGC
TAAGTAAGTAAGACGTCGAGCTCCCTATAGTGAGTCGTATTACACTGGCCGTCGTTTTAC
AACGTCGTGACTGGGAAAACACCGGTGAGCTCTAAGTAAGTAACGGCCGCCACCGCGGTG
GAGCTTTGGACTTCTTCGCCAGAGGTTTGGTCAAGTCTCCAATCAAGGTTGTGCGCTTGT
CTACCTTGCCAGAAATTTACGAAAAGATGGAAAAGGGTCAAATCGTTGGTAGATACGTTG
TTGACACTTCTAAATAAGCGAATTTCTTATGATTTATGATTTTTATTATTAAATAAGTTA
TAAAAAAAATAAGTGATACAAATTTTAAAGTGACTCTTAGGTTTTAAACGAAAATTCT
TGTTCTTGAGTAACCTCTTTCCTGTAGGTGAGGTTGCTTTCTCAGGTATAGCATGAGGTCG
CTCTTATTGACCACACCTCTACCGGCATGCCGAGCAAATGCCTGCAAATCGCTCCCCATT
TCACCCAATTGTAGATATGCTAACTCCAGCAATGAGTTGATGAATCTCGGTGTGTATTTT
ATGTCCTCAGAGGACAATACCTGTTGTAATCGTTCTTCCACACGGATCCGCATCAGGCGA
AATTGTAAACGTTAATATTTTGTAAATTCGCGTTAAATATTTGTAAATCAGCTCATT
TTTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAGAATAGACCGAGAT
AGGGTTGAGTGTTGTTCCAGTTTGGAAACAAGAGTCCACTATTAAAGAACGTGGACTCCAA
CGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACCCCTA
ATCAAGTTTTTTGGGGTTCGAGGTGCCGTAAAGCACTAAATCGGAACCCCTAAAGGGAGCCC
CCGATTTAGAGCTTGACGGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGGAAGAAAGC
GAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTACGCTGCGCGTAACCAACCAC
ACCCGCCGCGCTTAATGCGCCGCTACAGGGCGGTCCCATTCGCCATTCACTGCA

FIGURE 95D

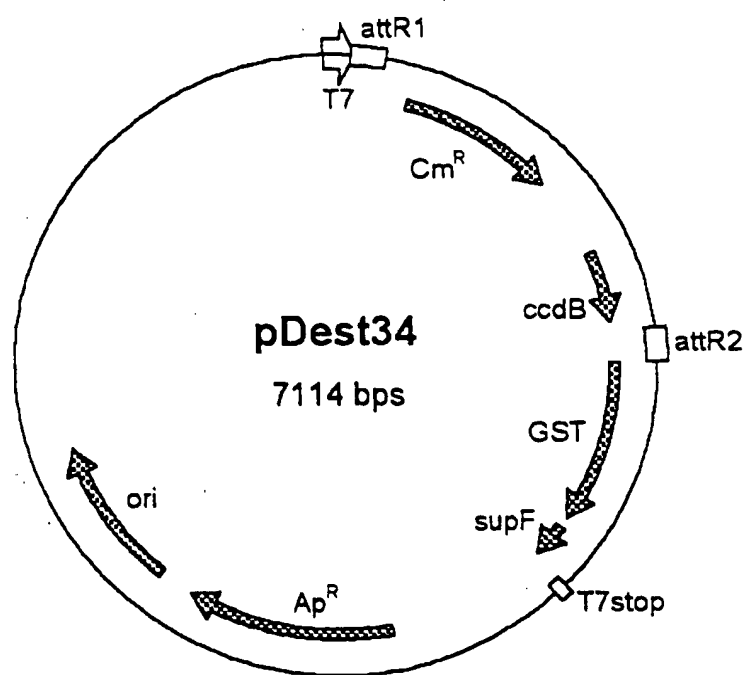


FIGURE 96A

pDEST34 7114 bp

<u>Location (Base Nos.)</u>	<u>Gene Encoded</u>
195..71	attR1
304..963	CmR
1305..1610	ccdB
1651..1775	attR2
1780..2472	GST
2675..2720	T7stop
3334..4194	ampR
4343..4982	ori

ATCGAGATCTCGATCCCGCGAAATTAATACGACTCACTATAGGGAGACCACAACGGTTTC
 CCTCTAGATCACAAGTTTGTACAAAAAGCTGAACGAGAAACGTAAAATGATATAAATAT
 CAATATATTAAATTAGATTTTGCATAAAAAACAGACTACATAACTGTAAAAACACAACA
 TATCCAGTCACTATGGCGGCCCGCATTAGGCACCCAGGCTTTACACTTTATGCTTCCGGC
 TCGTATAATGTGTGGATTTTGTAGTTAGGATCCGGCGAGATTTTCAGGAGCTAAGGAAGCT
 AAAATGGAGAAAAAATCACTGGATATACCACCGTTGATATATCCCAATGGCATCGTAAA
 GAACATTTTGAGGCATTTCACTCAGTTGCTCAATGTACCTATAACCAGACCGTTTCAGCTG
 GATATTACGGCCTTTTTAAAGACCGTAAAGAAAAATAAGCACAAAGTTTTATCCGGCCTTT
 ATTCACATTCTTGCCCGCCTGATGAATGCTCATCCGGAATTCGGTATGGCAATGAAAGAC
 GGTGAGCTGGTGATATGGGATAGTGTTCACCCCTTGTACACCGTTTTCCATGAGCAAAC
 GAAACGTTTTTCATCGCTCTGGAGTGAATACCACGACGATTTCCGGCAGTTTCTACACATA
 TATTCGCAAGATGTGGCGTGTACGGTGAAAAACCTGGCCTATTTCCCTAAAGGGTTTTATT
 GAGAAATATGTTTTTCGTCTCAGCCAATCCCTGGGTGAGTTTCACCAAGTTTGTATTTAAAC
 GTGGCCAATATGGACAACTTCTTCGCCCCCGTTTTACCATGGGCAAATATTATACGCAA
 GCGGACAAGGTGCTGATGCCGCTGGCGATTTCAGGTTTCATCATGCCGTCTGTGATGGCTTC
 CATGTCCGGCAGAATGCTTAATGAATTACAACAGTACTGCGATGAGTGGCAGGGCGGGGCG
 TAAACCGGTGGATCCGGCTTACTAAAAGCCAGATAACAGTATGCGTATTTGCGCGCTGAT
 TTTTGCGGTATAAGAATATATACTGATATGTATACCCGAAGTATGTCAAAAAGAGGTGTG
 CTATGAAGCAGCGTATTACAGTGACAGTTGACAGCGACAGCTATCAGTTGCTCAAGGCAT
 ATATGATGTCAATATCTCCGGTCTGGTAAGCACAAACCATGCAGAATGAAGCCCGTCGTCT
 GCGTGCCGAACGCTGGAAAGCGGAAATCAGGAAGGGATGGCTGAGGTGCGCCCGGTTTTAT
 TGAAATGAACGGCTCTTTTGTGCTGACGAGAACAGGGACTGGTGAAATGCAGTTTAAAGTTT
 ACACCTATAAAAGAGAGAGCCGTTATCGTCTGTTTGTGGATGTACAGAGTGATATTATTG
 ACACGCCCCGGGCGACGGATGGTGATCCCCCTGGCCAGTGCACGTCTGCTGTGATGATAAG
 TCTCCCGTGAACTTTACCCGGTGGTGATATCGGGGATGAAAGCTGGCGCATGATGACCA
 CCGATATGGCCAGTGTGCGGCTCTCCGTTATCGGGGAAGAAGTGGCTGATCTCAGCCACC
 GCGAAAAATGACATCAAAAACGCCATTAACTGATGTTCTGGGGAATATAAATGTGAGGCT
 CCCTTATACACAGCCAGTCTGCAGGTGACCATAGTGAAGTGGATATGTTGTGTTTTACAG
 TATTATGTAGTCTGTTTTTTATGCAAAATCTAATTTAATATATTGATATTTATATCATTT
 TACGTTTCTCGTTAGCTTTCTTGTACAAAGTGGTGATTATGTCCCCTATACTAGGTTAT
 TGGAAAAATTAAGGGCCTTGTGCAACCCACTCGACTTCTTTTGGAAATATCTTGAAGAAAA
 TATGAAGAGCATTTGTATGAGCGCGATGAAGGTGATAAATGGCGAAACAAAAAGTTTGAA
 TTGGGTTTGGAGTTTCCCAATCTTCTTATTATATTGATGGTGATGTTAAATTAACACAG
 TCTATGGCCATCATACGTTATATAGCTGACAAGCACAAACATGTTGGGTGGTTGTCCAAAA
 GAGCGTGACAGATTTCAATGCTTGAAGGAGCGGTTTTGGATATTAGATACGGTGTTTCG
 AGAATTGCATATAGTAAAGACTTTGAAACTCTCAAAGTTGATTTTCTTAGCAAGCTACCT
 GAAATGCTGAAAATGTTTGAAGATCGTTTATGTATATAAAACATATTTAAATGGTGATCAT
 GTAACCATCTGACTTCATGTTGTATGACGCTCTTGATGTTGTTTTATACATGGACCCA
 ATGTGCCCTGGATGCGTTCCCAAATTAGTTTGTGTTTTAAAAAACGTATTGAAGCTATCCCA
 CAAATTGATAAGTACTTGAAATCCAGCAAGTATATAGCATGGCCTTTGCAGGGCTGGCAA
 GCCACGTTTGGTGGTGGCGACCATCTCCAAAATCGGATCTGGTTCGCGCTCCATGGGGA
 TCCGGCTGCTAACAAAGCCCCGAAAGGAAGCTGAGTTGGCTGCTGCCACCGCTGAGCGCTT
 CCCGATAAGGGAGCAGGCCAGTAAAGCATTACCCGTGGTGGGGTTCCCGAGCGGCCAAA
 GGGAGCAGACTCTAAATCTGCCGTCTGCACTTCGAAGGTTTGAATCCTTCCCCCACCAC
 CATCACTTTCAAAGTGAATTCGCTGAGCAATAACTAGCATAACCCCTTGGGGCCTCTAA-

FIGURE 96B

ACGGGTCTTGAGGGGTTTTTTTGGCTGAAAGGAGGAACTATATCCGGATATCCACAGGACGG
GTGTGGTTCGCCATGATCGCGTAGTCGATAGTGGCTCCAAGTAGCGAAGCGAGCAGGACTG
GGCGGCGGCCAAAGCGGTTCGGACAGTGTCTCCGAGAACGGGTGCGCATAGAAATTGCATCA
ACGCATATAGCGCTAGCAGCACGCCATAGTGACTGGCGATGCTGTGCGAATGGACGATAT
CCCCAAGAGGCCCCGGCAGTACCGGCATAACCAAGCCTATGCCTACAGCATCCAGGGTGA
CGGTGCCGAGGATGACGATGAGCGCATTGTTAGATTTCATACACGGTGCCTGACTGCGTT
AGCAATTTAACTGTGATAAACTACCGCATTAAAGCTTATCGATGATAAGCTGTCAAACAT
GAGAATTCCTTGAAGACGAAAGGGCCTCGTGATACGCCCTATTTTTATAGGTTAATGTCATG
ATAATAATGGTTTCTTAGACGTCAGGTGGCACTTTTCGGGGAAATGTGCGCGGAACCCCT
ATTTGTTTTATTTTTCTAAATACATTCAAATATGTATCCGCTCATGAGACAATAACCCCTGA
TAAATGCTTCAATAATATTGAAAAAGGAAGAGTATGAGTATTCAACATTTCCGTGTGCGC
CTTATTCCTTTTTTTGCGGCATTTTGCCTTCTGTTTTTGGCTCACCCAGAAACGCTGGTG
AAAGTAAAGATGCTGAAGATCAGTTGGGTGCACGAGTGGGTACATCGAACTGGATCTC
AACAGCGGTAAGATCCTTGAGAGTTTTTCGCCCGAAGAACGTTTTCCAATGATGAGCACT
TTTAAAGTTCTGCTATGTGGCGCGGTATTATCCCGTGTGACGCCGGGCAAGAGCAACTC
GGTCGCCGCATACACTATTCTCAGAATGACTTGGTTGAGTACTCACCAGTCACAGAAAAG
CATCTTACGGATGGCATGACAGTAAGAGAATTATGCAGTGCTGCCATAACCATGAGTGAT
AACACTGCGGCCAACTTACTTCTGACAACGATCGGAGGACCGAAGGAGCTAACCGCTTTT
TTGCACAACATGGGGGATCATGTAACCTCGCCTTGATCGTTGGGAACCGGAGCTGAATGA
GCCATACCAAACGACGAGCGTGACACCAGTATGCTGCAGCAATGGCAACAACGTTGCGC
AACTATTAACTGGCGAACTACTTACTCTAGCTTCCCGCAACAATTAATAGACTGGATG
GAGGCGGATAAAGTTGACAGGACCACTTCTGCGCTCGGCCCTTCCGGCTGGCTGGTTTATT
GCTGATAAATCTGGAGCCGGTGAGCGTGGGTCTCGCGGTATCATTGCAGCACTGGGGCCA
GATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAGGCAACTATGGAT
GAACGAAATAGACAGATCGCTGAGATAGGTGCCTCACTGATTAAGCATTGGTAACTGTCA
GACCAAGTTTACTCATATATACTTTAGATTGATTTAAACTTCATTTTTAATTTAAAGG
ATCTAGGTGAAGATCCTTTTTGATAATCTCATGACCAAAATCCCTTAACGTGAGTTTTCG
TTCCACTGAGCGTCAGACCCCGTAGAAAAAGATCAAAGGATCTTCTTGAGATCCTTTTTTT
CTGCGCGTAATCTGCTGCTTGCAACAAAAAACACCGCTACCAGCGGTGGTTTTGTTTG
CCGGATCAAGAGCTACCAACTCTTTTTCCGAAGGTAAGTGGCTTCAGCAGAGCGCAGATA
CCAAATACTGTCTTCTAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCA
CCGCCTACATACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGATAAG
TCGTGTCTTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTGCGGC
TGAACGGGGGGGTTCTGTCACACAGCCAGCTTGGAGCGAAGCGACCTACACCGAACTGAGA
TACCTACAGCGGTGAGCTATGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGG
TATCCGGTAAGCGGCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAC
GCCTGGTATCTTTATAGTCTGTGCGGTTTTCGCCACCTCTGACTTGAGCGTTCGATTTTTG
TGATGCTCGTCAGGGGGGCGGAGCCTATGGAAAAACGCCAGCAACGCGGCCTTTTTACGG
TTCCTGGCCTTTTTGCTGGCCTTTTTGCTCACATGTTCTTTCCTGCGTTATCCCCTGATTCT
GTGGATAACCGTATTACCGCCTTTGAGTGAGCTGATACCGCTCGCCGCGAGCCGAACGACC
GAGCGCAGCGAGTCAGTGAGCGAGGAAGCGGAAGCGCCTGATGCGGTATTTTTCTCCTT
ACGCATCTGTGCGGTATTTTACACCGCATATATGGTGCACCTCTCAGTACAATCTGCTCTG
ATGCCGCATAGTTAAGCCAGTATACACTCCGCTATCGCTACGTGACTGGGTCTGGCTGC
GCCCCGACACCCGCCAACACCCGCTGACGCGCCCTGACGGGCTTGTCTGCTCCCGGCATC
CGCTTACAGACAAGCTGTGACCGTCTCCGGGAGCTGCATGTGTGAGAGGTTTTACCGTC
ATCACCGAAACGCGCGAGGCGAGCTGCGGTAAAGCTCATCAGCGTGGTTCGTGAAGCGATT
ACAGATGTCTGCCTGTTTATCCGCGTCCAGCTCGTTGAGTTTTCTCCAGAAGCGTTAATGT
CTGGCTTCTGATAAAGCGGGCCATGTTAAGGGCGGTTTTTTTCTGTTTGGTCACTGATGC
CTCCGTGTAAGGGGGATTTCTGTTTATGGGGGTAATGATACCGATGAAACGAGAGAGGAT
GCTCACGATACGGGTTACTGATGATGAACATGCCCGGTTACTGGAACGTTGTGAGGGTAA
ACAACTGGCGGTATGGATGCGGCGGGACAGAGAAAAATCACTCAGGGTCAATGCCAGCG
CTTCGTTAATACAGATGTAGGTGTTCCACAGGGTAGCCAGCAGCATCCTGCGATGCAGAT
CCGGAACATAATGGTGCAGGGCGCTGACTTCCGCGTTTCCAGACTTTACGAAACACGGAA
ACCGAAGACCATTCATGTTGTTGCTCAGGTGCAGACGTTTTGACAGCAGCAGTCGCTTCA
CGTTGCTCGCGTATCGGTGATTCTTCTGCTAACAGTAAGGCAACCCCGCCAGCCTAG
CCGGGTCTTCAACGACAGGAGCACGATCATGCGCACCCGTGGCCAGGACCCAACGCTGCC
CGAGATGCGCCGCTGCGGCTGCTGGAGATGGCGGACGCGATGGATATGTTCTGCCAAGG
GTTGGTTTTGCGCATTACAGTTCTCCGCAAGAATTGATTGGCTCCAATTCTTGGAGTGGT-

FIGURE 96C

GAATCCGTTAGCGAGGTGCCGCCGGCTTCCATTACAGGTCGAGGTGGCCCCGGCTCCATGCA
CCGCGACGCAACGCGGGGAGGCAGACAAGGTATAGGGCGGCGCCTACAATCCATGCCAAC
CCGTTCCATGTGCTCGCCGAGGCGGCATAAATCGCCGTGACGATCAGCGGTCCAGTGATC
GAAGTTAGGCTGGTAAGAGCCGCGAGCGATCCTTGAAGCTGTCCCTGATGGTCGTCTCT
ACCTGCCTGGACAGCATGGCCTGCAACGCGGGCATCCCGATGCCGCCGGAAGCGAGAAGA
ATCATAATGGGGAAGGCCATCCAGCCTCGCGTCGCGAACGCCAGCAAGACGTAGCCCAGC
GCGTCGGCCGCCATGCCGGCGATAATGGCCTGCTTCTCGCCGAAACGTTTGGTGGCGGGA
CCAGTGACGAAGGCTTGAGCGAGGGCGTGCAAGATTCCGAATACCGCAAGCGACAGGCCG
ATCATCGTCGCGCTCCAGCGAAAGCGGTCTCTCGCCGAAAATGACCCAGAGCGCTGCCGGC
ACCTGTCCTACGAGTTGCATGATAAAGAAGACAGTCATAAGTGCGGCGACGATAGTCATG
CCCCGCGCCACCGGAAGGAGCTGACTGGGTTGAAGGCTCTCAAGGGCATCGGTTCGATCG
ACGCTCTCCCTTATGCGACTCCTGCATTAGGAAGCAGCCAGTAGTAGGTTGAGGCCGTT
GAGCACCGCCGCCGCAAGGAATGGTGATGCAAGGAGATGGCGCCCAACAGTCCCCCGGC
CACGGGGCCTGCCACCATACCCACGCCGAAACAAGCGCTCATGAGCCCGAAGTGGCGAGC
CCGATCTTCCCATCGGTGATGTGGCGGATATAGGCGCCAGCAACCGCACCTGTGGCGCC
GGTGATGCCGGCCACGATGCGTCCGGCGTAGAGG

FIGURE 96A

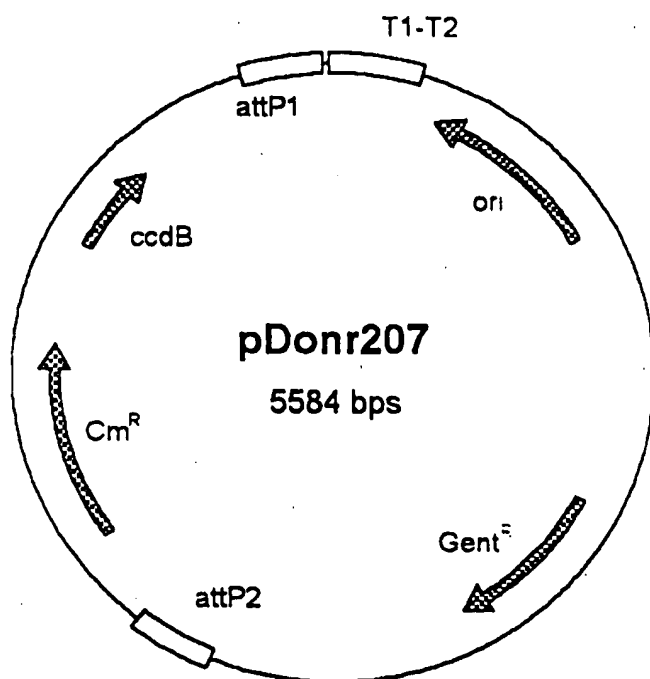


FIGURE 97A

pDONR207

5584 bp

GCGAGAGTAGGGAACTGCCAGGCATCAAATAAAACGAAAGGCTCAGTCGGAAGACTGGGC
CTTTCGTTTTATCTGTTGTTTGTTCGGTGAACGCTCTCCTGAGTAGGACAAATCCGCCGGG
AGCGGATTTGAACGTTGTGAAGCAACGGCCCCGAGGGTGGCGGGCAGGACGCCCGCCATA
AACTGCCAGGCATCAAACCTAAGCAGAAGGCCATCCTGACGGATGGCCTTTTTGCGTTTCT
ACAAACTCTTCCTGGCTAGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGA
AAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCTTGCTG
GCGTTTTTTCATAGGCTCCGCCCCCTGACGAGCATCACAAAATCGACGCTCAAGTCAG
AGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTTCCCCCTGGAAGCTCCCTC
GTGCGCTCTCCTGTTCCGACCCCTGCCGCTTACCGGATACCTGTCCGCTTTCTCCCTTCG
GGAAGCGTGGCGCTTTCTCATAGCTCAGCTGTAGGTATCTCAGTTCGGGTGATAGGTCGTT
CGTCCAAAGCTGGGCTGTGTGCACGAACCCCGTTTCAGCCCGACCGCTGCGCTTATCC
GGTAACATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCC
ACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTCTTGAAGTGG
TGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCA
GTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGC
GGTGGTTTTTTTGTGTTGCAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGAT
CCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAAACGAAAACCTCAGTTAAGGGATT
TTGGTCATGAGTTGCGCGTCCCGTCAAGTCAGCGTAATGCTCTGCCAGTGTACAACC
AATTAACCAATTCTGATTAGAAAACTCATCGAGCATCAAATGAACTGCAATTTATTCA
TATCAGGATTATCAATACCATATTTTGAAGAACCGTTTTCTGTAATGAAGGAGAAAACT
CACCGAGGCAGTTCATAGGATGGCAAGATCCTGGTATCGGTCTGCGATTCCGACTCGTC
CAACATCAATACAACCTATTAGTAGCCAACCACTAGAACTATAGCTAGAGTCTTGGGCGA
ACAAACGATGCTCGCCTTCAGAAAACCGAGGATGCGAACCCTTATCCGGGGTCAGCA
CCACCGGCAAGCGCCGCGACGGCCGAGGTCTTCGATCTCCTGAAGCCAGGGCAGATCCG
TGCACAGCACCTTGCCGTAGAAGAACAGCAAGGCCCAATGCCTGACGATGCGTGGAGA
CCGAAACCTTGCGCTCGTTGCGCAGCCAGGACAGAAATGCCTCGACTTCGCTGCTGCCCA
AGGTTGCCGGGTGACGCACACCGTGGAACGGATGAAGGCACGAACCCAGTTGACATAAG
CCTGTTGCGTTCGTAACTGTAATGCAAGTAGCGTATGCGCTCACGCAACTGGTCCAGAA
CCTTGACCGAAGCGCAGCGGTGGTAACGGCGCAGTGGCGGTTTTTCATGGCTTGTTATGACT
GTTTTTTTGTACAGTCTATGCCTCGGGCATCCAAGCAGCAAGCGGTTACGCCGTGGGT
GATGTTTGATGTTATGGAGCAGCAACGATGTTACGCAGCAGCAACGATGTTACGCAGCAG
GGCTCGGCTCTCTATACAAAGTTAGGTGGCTCAAGTATGGGCATCATTGACATGTAGG
CTCGGCCCTGACCAAGTCAAATCCATGCGGGCTGCTCTTGATCTTTTCGGTTCGTGAGTTC
GGAGACGTAGCCACCTACTCCAACATCAGCCGGACTCCGATTACCTCGGGAACCTTGCTC
CGTAGTAAGACATTTCATCGCGCTTGCTGCCTTCGACCAAGAAGCGGTTGTTGGCGCTCTC
GCGGCTTACGTTCTGCCCAGGTTTGAGCAGCCGCGTAGTGAGATCTATATCTATGATCTC
GCAGTCTCCGGCGAGCACCGGAGGCAGGGCATTGCCACCGCGCTCATCAATCTCCTCAAG
CATGAGGCCAACGCGCTTGGTGCTTATGTGATCTACGTGCAAGCAGATTACGGTGACGAT
CCCGCAGTGGCTCTCTATACAAAGTTGGGCATACGGGAAGAAGTGATGCACTTTGATATC
GACCAAGTACCGCCACCTAACAATTCGTTCAAGCCGAGATCGGCTTCCCGGCCTAATTT
CCCCCTCGTCAAAAATAAGGTTATCAAGTGAGAAATCACCATGAGTGACGACTGAATCCGG
TGAGAAATGGCAAAAGTTTATGCATTTCTTTCCAGACTTGTTCAACAGGCCAGCCATTACG
CTCGTCATCAAATCACTCGCATCAACCAAACCGTTATTCATTCTGTGATTGCGCCTGAGC
GAGACGAAATACGCGATCGCTGTTAAAGGACAATTACAAACAGGAATCGAATGCAACCG
GCGCAGGAACACTGCCAGCGCATCAACAATATTTTACCTGAATCAGGATATTCTTCTAA
TACCTGGAATGCTGTTTTTCCGGGGATCGCAGTGGTGAGTAACCATGCATCATCAGGAGT
ACGGATAAAATGCTTGATGGTTCGGAAGAGGCATAAAATCCGTCAGCCAGTTTGTCTGAC
CATCTCATCTGTAAACATCATTGGCAACGCTACCTTTGCCATGTTTCAGAAACAACCTCTGG
CGCATCGGGCTTCCCATACAAGCGATAGATTGTGCGACCTGATTGCCCGACATTATCGCG
AGCCCATTTATACCCATATAAATCAGCATCCATGTTGGAATTTAATCGCGGCCCTCGACGT
TTCCCGTTGAATATGGCTCATAACACCCCTGTATTACTGTTTATGTAAGCAGACAGTTT
TATTGTTTCATGATGATATATTTTTATCTTGTGCAATGTAACATCAGAGATTTTGAGACAC
GGCCAGAGCTGCAGCTGGATGGCAATAATGATTTTATTTTGACTGATAGTGACCTGTT
CGTTGCAACAAATGATAAGCAATGCTTTCTTATAATGCCAATTTGTACAAGAAAGCTG
AACGAGAAACGTAAAATGATATAAATATCAATATATTAATTAGATTTTGCATAAAAAAC
AGACTACATAATACTGTAAAACACAACATATCCAGTCACTATGAATCAACTACTTAGATG-

FIGURE 97B

GTATTAGTGACCTGTAGTCGACTAAGTTGGCAGCATCACCCGACGCACTTTGCGCCGAAT
AAATACCTGTGACGGAAGATCACTTCGAGAATAAATAAATCCTGGTGTCCCTGTTGATA
CCGGGAAGCCCTGGGCCAACTTTGGCGAAAATGAGACGTTGATCGGCACGTAAGAGGTTT
CAACTTTTACCATAATGAAATAAGATCACTACCGGGCGTATTTTTTTGAGTTATCGAGATT
TTCAGGAGCTAAGGAAGCTAAAATGGAGAAAAAATCACTGGATATACCACCGTTGATAT
ATCCCAATGGCATCGTAAAGAACATTTTGAGGCATTTTCACTCAGTTGCTCAATGTACCTA
TAACCAGACCGTTTCACTGGATATTACGGCCTTTTTTAAAGACCGTAAAGAAAAATAAGCA
CAAGTTTTATCCGGCCTTTATTACATTCTTGCCCGCCTGATGAATGCTCATCCGGAATT
CCGTATGGCAATGAAAGACGGTGAGCTGGTGATATGGGATAGTGTTCAACCTTGTTACAC
CGTTTTTCCATGAGCAAACCTGAAACGTTTTTTCATCGCTCTGGAGTGAATACCACGACGATTT
CCGGCAGTTTTCTACACATATATTTCGAAGATGTGGCGTGTTACGGTGAAAACCTGGCCTA
TTTTCCCTAAAGGGTTTTATTGAGAATATGTTTTTTCGTCTCAGCCAATCCCTGGGTGAGTTT
CACCAGTTTTGATTTAAACGTGGCCAATATGGACAACCTTCTTCCCCCGTTTTTACCAT
GGGCAAATATTATACGCAAGGCGACAGGTGCTGATGCCGCTGGCGATTTCAGGTTTCATCA
TGCCGTCTGTGATGGCTTCCATGTCGGCAGAATGCTTAATGAATTACAACAGTACTGCCA
TGAGTGGCAGGGCGGGGCGTAATCGCGTGGATCCGGCTTACTAAAAGCCAGATAACAGTA
TGCGTATTTGCGCGCTGATTTTTGCGGTATAAGAATATATACTGATATGTATACCCGAAG
TATGTCAAAAAGAGGTGTGCTATGAAGCAGCGTATTACAGTGACAGTTGACAGCGACAGC
TATCAGTTGCTCAAGGCATATATGATGTCAATATCTCCGGTCTGGTAAGCACAACCATGC
AGAATGAAGCCCGTCTGCTGCGTGCCGAACGCTGGAAAGCGGAAAATCAGGAAGGGATGG
CTGAGGTGCGCCGGTTTTATTGAAATGAACGGCTCTTTTGCTGACGAGAACAGGGACTGGT
GAAATGCAGTTTAAAGGTTTACACCTATAAAAGAGAGAGCCGTTATCGTCTGTTTGTGGAT
GTACAGAGTGATATTATTGACACGCCCCGGGCGACGGATGGTGATCCCCCTGGCCAGTGCA
CGTCTGCTGTCAGATAAAGTCTCCCGTGAACCTTTACCCGGTGGTGATATCGGGGATGAA
AGCTGGCGCATGATGACCACCGATATGGCCAGTGTGCCGGTCTCCGTTATCGGGGAAGAA
GTGGCTGATCTCAGCCACCCGCGAAAATGACATCAAAAACGCCATTAACTGATGTTCTGG
GGAATATAAATGTCAGGCTCCCTTATACACAGCCAGTCTGCAGGTCGATACAGTAGAAAT
TACAGAACTTTTATCACGTTTAGTAAGTATAGAGGCTGAAAATCCAGATGAAGCCGAACG
ACTTGTAAGAGAAAAGTATAAGAGTTGTGAAATTGTTCTTGATGCAGATGATTTTCAGGA
CTATGACACTAGCGTATATGAATAGGTAGATGTTTTTATTTTGTACACAAAAAAGAGGC
TCGCACCTCTTTTTCTTATTTCTTTTTATGATTTAATACGGCATTGAGGACAATAGCGAG
TAGGCTGGATACGACGATTCCGTTTTGAGAAGAACATTTGGAAGGCTGTCCGTCGACTAAG
TTGGCAGCATCACCCGAAGAACATTTGGAAGGCTGTCCGTCGACTACAGGTCACTAATAC
CATCTAAGTAGTTGATTCATAGTGAAGTGGATATGTTGTGTTTTTACAGTATTATGTAGTCT
GTTTTTTATGCAAAATCTAATTTAATATATTGATATTTATATCATTTTACGTTTCTCGTT
CAGCTTTTTTGTACAAAGTTGGCATTATAAAAAAGCATTTGCTCATCAATTTGTTGCAACG
AACAGGTCACTATCAGTCAAATAAATCATTATTTGGGGCCCGAGATCCATGCTAGCGT
TAAC

FIGURE 97C

pMAB85

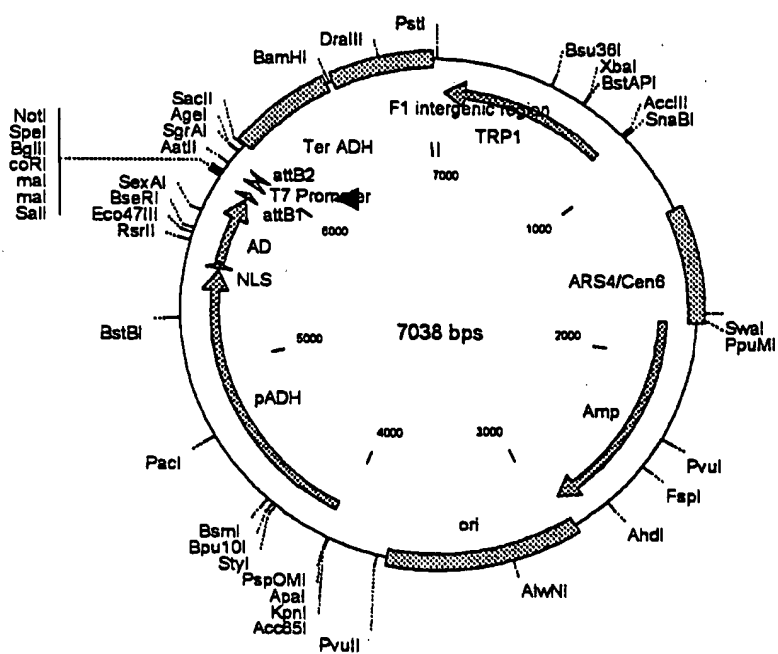


FIGURE 98A

pMAB85 7038 bp

GCCTTACGCATCTGTGCGGTATTTTCACACCCGAGGCAAGTGCACAAACAATACTTAAATA
AATACTACTCAGTAATAACCTATTTCTTAGCATTTTGTGACGAAATTTGCTATTTTGTAG
AGTCTTTTACACCATTTGTCTCCACACCTCCGCTTACATCAACACCAATAACGCCATTTA
ATCTAAGCGCATCACCAACATTTTCTGGCGTCAGTCCACCAGCTAACATAAAATGTAAGC
TTTCGGGGCTCTCTTGCCTTCCAACCCAGTCAGAAATCGAGTTCCAATCCAAAAGTTCAC
CTGTCCCACCTGCTTCTGAATCAAACAAGGGAATAAACGAATGAGGTTTCTGTGAAGCTG
CACTGAGTAGTATGTTGTCAGTCTTTTGGAAATACGAGTCTTTTAATAACTGCGCAAACCGA
GGAACCTCTTGGTATTCTTGCCACGACTCATCTCCATGCAGTTGGACGATATCAATGCCGT
AATCATTGGACAGCCAAAACATCCTCCTTAGGTTGATTACGAAACACGCCAACCAAGT
ATTTCCGGAGTGCCTGAACTATTTTATATGCTTTTACAAGACTTGAAATTTTCCTTGCAA
TAACCGGGTCAATTGTTCTCTTCTATTGCGGCACACATATAATACCCAGCAAGTCAGCAT
CGGAATCTAGAGCACATTCTGCGGCCCTCTGTGCTCTGCAAGCCGCAAACCTTTCACCAATG
GACCAGAACTACCTGTGAAATTAATAACAGACATACTCCAAGCTGCCTTTGTGTGCTTAA
TCACGTATACTCACGTGCTCAATAGTCACCAATGCCCTCCCTCTTGGCCCTCTCCTTTTC
TTTTTTCGACCGAATTAATTCTTAATCGGCAAAAAAAGAAAAGCTCCGGATCAAGATTGT
ACGTAAGGTGACAAAGCTATTTTCAATAAAGAATATCTTCCACTACTGCCATCTGGCGTC
ATAACTGCAAAGTACACATATATTACGATGCTGTCTATTAAATGCTTCCTATATTATATA
TATAGTAATGTCGTTTATGGTGCACTCTCAGTACAATCTGCTCTGATGCCGCATAGTTAA
GCCAGCCCCGACACCCGCCAACACCCGCTGACGCGCCCTGACGGGCTTGTCTGCTCCCGG
CATCCGCTTACAGACAAGCTGTGACCGTCTCCGGGAGCTGCATGTGTGAGAGGTTTTAC
CGTCATCACCGAAACGCGCGAGACGAAAGGGCCTCGTGATACGCCCTATTTTTATAGGTTA
ATGTCTATGATAATAATGGTTTCTTAGGACGGATCGCTTGCCTGTAACCTACACGCGCCTC
GTATCTTTTAAATGATGGAATAATTTGGGAATTTACTCTGTGTTTATTTATTTTATGTTT
TGTATTTGGATTTTAGAAAGTAAATAAAGAAGGTAGAAGAGTTACGGAATGAAGAAAAA
AAATAAACAAAGGTTTTAAAAATTTCAACAAAAGCGTACTTTACATATATATTTATTAG
ACAAGAAAAGCAGATTAAATAGATATACATTGATTAACGATAAGTAAATGTAAATCA
CAGGATTTTCTGTGTGCTCTTCTACACAGACAAGATGAAACAATTCGGCATTAAATACCT
GAGAGCAGGAAGAGCAAGATAAAAGGTAGTATTTGTTGGCGATCCCCCTAGAGTCTTTTA
CTATCTTCGGAAAACAAAACCTATTTTTTCTTTAATTTCTTTTTTACTTTCTATTTTAA
TTTATATATTTATATTTAAAAAATTTAAATTATAATTATTTTATAGCACGTGATGAAAAG
GACCCAGGTGGCACTTTTCCGGGAAATGTGCGCGGAACCCCTATTTGTTTATTTTCTAA
ATACATTCAAATATGTATCCGCTCATGAGACAATAACCCCTGATAAATGCTTCAATAATAT
TGAAAAGGAAGAGTATGAGTATTCAACATTTCCGTGTCGCCCTTATTCCTTTTTTGCG
GCATTTTGCCTTCTGTTTTTTGCTCACCCAGAAACGCTGGTGAAAGTAAAGATGCTGAA
GATCAGTTGGGTGCACGAGTGGGTACACTCGAAGTGGATCTCAACAGCGGTAAGATCCTT
GAGAGTTTTCGCCCCGAAGAACGTTTTTCCAATGATGAGCACTTTTAAAGTTCTGCTATGT
GGCGCGGTATTATCCCGTATTGACGCGCGGCAAGAGCAACTCGGTGCGCGCATACACTAT
TCTCAGAATGACTTGGTTGAGTACTCACAGTCACAGAAAAGCATCTTACGGATGGCATG
ACAGTAAGAGAATTATGCAGTGTGCCATAACCATGAGTGATAACACTGCGGCCAACTTA
CTTCTGACAACGATCGGAGGACCGAAGGAGCTAACCGCTTTTTTTCACAACATGGGGGAT
CATGTAACCTGCCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAACGACGAG
CGTGACACCACGATGCCTGTAGCAATGGCAACAACGTTGCGCAAACTATTAAGTGGCGAA
CTACTTACTCTAGCTTCCCGCAACAATTGAATAGACTGGATGGAGGCGGATAAAGTTGCA
GGACCACTTCTGCGCTCGGCCCTTCCGGCTGGCTGGTTTATTGCTGATAAATCTGGAGCC
GGTGAGCGTGGGTCTCGCGGTATCATTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGT
ATCGTAGTTATCTACACGACGGGCAGTCAGGCAACTATGGATGAACGAAATAGACAGATC
GCTGAGATAGGTGCCTCACTGATTAAGCATTGGTAACTGTGACACCAAGTTTACTCATAT
ATACTTTAGATTGATTTAAAACTTCATTTTTAATTTAAAGGATCTAGGTGAAGATCCTT
TTTGATAATCTCATGACCAAAATCCCTTAACGTGAGTTTTCGTTCCTGAGCGTCAGAC
CCCGTAGAAAAGATCAAAGGATCTTCTTGAGATCCTTTTTTTCTGCGCGTAATCTGCTGC
TTGCAACAAAAAACACCGCTACCAGCGGTGGTTTGTGTTGCCGGATCAAGAGCTACCA
ACTCTTTTTCCGAAGGTAACTGGCTTCAGCAGAGCGCAGATACCAAATACTGTCTCTTA
GTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCT
CTGCTAATCCTGTTACAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTTG
GACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTGCGGCTGAACGGGGGGTTCGTGC-

Figure 98B

ACACAGCCCAGCTTGGAGCGAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCAT
TGAGAAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGCGGCAGG
GTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGGAACGCCTGGTATCTTTATAGT
CCTGTCGGGTTTTCGCCACCTCTGACTTGAGCGTCGATTTTTGTGATGCTCGTCAGGGGGG
CCGAGCCTATGGAAAAACGCCAGCAACGCGGCCTTTTTACGGTTCCTGGCCTTTTGCTGG
CCTTTTGCTCACATGTTCTTTCCTGCGTTATCCCCTGATTCTGTGGATAACCGTATTACC
GCCTTTGAGTGAGCTGATAACGCTCGCCGAGCCGAACGACCGAGCGCAGCGAGTCAGTG
AGCGAGGAAGCGGAAGAGCGCCCAATACGCAAACCGCCTCTCCCCGCGCGTTGGCCGATT
CATTAATGCAGCTGGCACGACAGGTTTCCCGACTGGAAAGCGGGCAGTGAGCGCAACGCA
ATTAATGTGAGTTACCTCACTCATTAGGCACCCAGGCTTTACACTTTATGCTTCCGGCT
CCTATGTTGTGTGGAATTGTGAGCGGATAACAATTTACACAGGAAACAGCTATGACCAT
GATTACGCCAAGCTCGGAATTAACCCTCACTAAAGGGAACAAAAGCTGGGTACCGGGCCC
CCCCTCGAGATCCGGGATCGAAGAAATGATGGTAAATGAAATAGGAAATCAAGGAGCATG
AAGGCAAAAGACAAATATAAGGGTTCGAACGAAAAATAAAGTGAAAAGTGTTGATATGATG
TATTTGGCTTTTGGCGCCGAAAAACAGAGTTTACGCAATTGCACAATCATGCTGACTCT
GTGGCGGACCCGCGCTCTTGCCGGCCCCGGCGATAACGCTGGGCGTGAGGCTGTGCCCGGC
GGAGTTTTTTGCGCCTGCATTTTCCAAGGTTTACCCTGCGCTAAGGGGCGAGATTGGAGA
AGCAATAAGAATGCCGTTGGGGTTGCGATGATGACGACCACGACAACTGGTGTCAATTAT
TTAAGTTGCGGAAAGAACCTGAGTGCAATTTGCAACATGAGTATACTAGAAGAATGAGCCA
AGACTTGCGAGACGCGAGTTTGCCGGTGGTGCGAACAATAGAGCGACCATGACCTTGAAG
GTGAGACGCGCATAACCGCTAGAGTACTTTGAAGAGGAAACAGCAATAGGGTTGCTACCA
GTATAAATAGACAGGTACATAACAACACTGGAAATGGTTGTCTGTTGAGTACGCTTTCAA
TTCATTTGGGTGTGCACTTTATTATGTTACAATATGGAAGGGAACCTTACACTTCTCCTA
TGCACATATATTAATTAAGTCCAATGCTAGTAGAGAAGGGGGGTAACACCCCTCCGCGC
TCTTTTCCGATTTTTTTCTAAACCGTGGAATATTTTCGGATATCCTTTTGTGTTTCCGGG
TGTACAATATGGACTTCTCTTTTCTGGCAACCAAACCCATACATCGGGATTCTCTATAAT
ACCTTCGTTGGTCTCCCTAACATGTAGGTGGCGGAGGGGAGATATACAATAGAACAGATA
CCAGACAAGACATAATGGGCTAAACAAGACTACACCAATTACACTGCCTCATTGATGGTG
GTACATAACGAACTAATACTGTAGCCCTAGACTTGATAGCCATCATCATATCGAAGTTTC
ACTACCCTTTTTCCATTTGCCATCTATTGAAGTAATAATAGGCGCATGCAACTTCTTTTC
TTTTTTTTTCTTTTCTCTCTCCCCGTTGTTGTCTCACCATATCCGCAATGACAAAAAAA
ATGATGGAAGACACTAAAGGAAAAAATTAACGACAAAGACAGCACCACAGATGTCGTTG
TTCCAGAGCTGATGAGGGGTATCTTCGAACACACGAAACTTTTTCTTCTTTCATTACAG
CACACTACTCTCTAATGAGCAACGGTATACGGCCTTCTTCCAGTTACTTGAATTTGAAA
TAAAAAAGGTTTGGCGCTTTGCTATCAAGTATAAATAGACCTGCAATTATTAATCTTTTG
TTTCTCGTCAATTGTTCTCGTTCCCTTTCTTCTTGTCTTTTTCTGCACAATATTTCA
AGCTATACCAAGCATACAATCAACTCCAAGCTTATGCCCAAGAAGAAGCGGAAGTCTCG
AGCGGCGCCAATTTTAATCAAAGTGGGAATATTGCTGATAGCTCATTGTCTTCACTTTC
ACTAACAGTAGCAACGGTCCGAACCTCATAACAACCTCAAACAAATTCTCAAGCGCTTCA
CAACCAATTGCCTCCTCTAACGTTTCATGATAACTTCATGAATAATGAAATCACGGCTAGT
AAAATTGATGATGGTAATAATTCAAAACCACTGTCACCTGGTTGGACGGACCAAACTGCG
TATAACGCGTTTGGAACTACTACAGGGATGTTTAAATACCACTACAATGGATGATGTATAT
AACTATCTATTTCGATGATGAAGATACCCCAACCAACCAAAAAAGAGGGTGGGTGATC
ACAAGTTTGTACAAAAAGCAGGCTTGTGACCCCGGGAATTGAGATCTACTAGTGGCGC
CGCACGCGTACCCAGCTTCTTGTACAAAGTGGTGACGTGAGCTCCCTATAGTGAGTCG
TATTACACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACACCGGTGAGCTCTAAGT
AAGTAACGGCCGCCACCGCGGTGGAGCTTTGGACTTCTTCGCCAGAGGTTTGGTCAAGTC
TCCAATCAAGGTTGTGCGCTTGTCTACCTTGCCAGAAATTTACGAAAAGATGGAAAAGGG
TCAAATCGTTGGTAGATACGTTGTTGACACTTCTAAATAAGCGAATTTCTTATGATTTAT
GATTTTTTATTATTAAATAAGTTATAAAAAAATAAGTGTATACAAATTTTAAAGTGAATC
TTAGGTTTTTAAACGAAATTTCTTGTCTTGAGTAACTCTTCTCTGATAGGTGAGTTGCT
TTCTCAGGTATAGCATGAGGTGCTCTTATTGACCACACCTCTACCGCATGCCGAGCAA
ATGCCTGCAAATCGCTCCCATTTACCCCAATTGTAGATATGCTAACTCCAGCAATGAGT
TGATGAATCTCGGTGTGATTTTTATGTCCTCAGAGGACAATACCTGTTGTAATCGTTCTT
CCACACGGATCCGCATCAGGCGAAATTTGTAAACGTTAATATTTGTAAAATTCGCGTTA
AATATTTGTAAATCAGCTCATTTTTTAACCAATAGGCCGAAATCGGCAAAATCCCTTAT
AAATCAAAAGAAATAGACCGAGATAGGGTTGAGTGTGTTCCAGTTTGGAAACAAGAGTCCA
CTATTAAAGAACGTGGACTCCAACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGC-

FIGURE 98C

CCACTACGTGAACCATCACCCCTAATCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTA
AATCGGAACCCTAAAGGGAGCCCCGATTTAGAGCTTGACGGGGAAAGCCGGCGAACGTG
GCGAGAAAGGAAGGGAAGAAAGCGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCG
GTCACGCTGCGCGTAACCACCACCCGCGCGCTTAATGCGCCGCTACAGGGCGCGTCC
CATTGCCCATTCACTGCA

FIGURE 98D

237/240

pMAB86

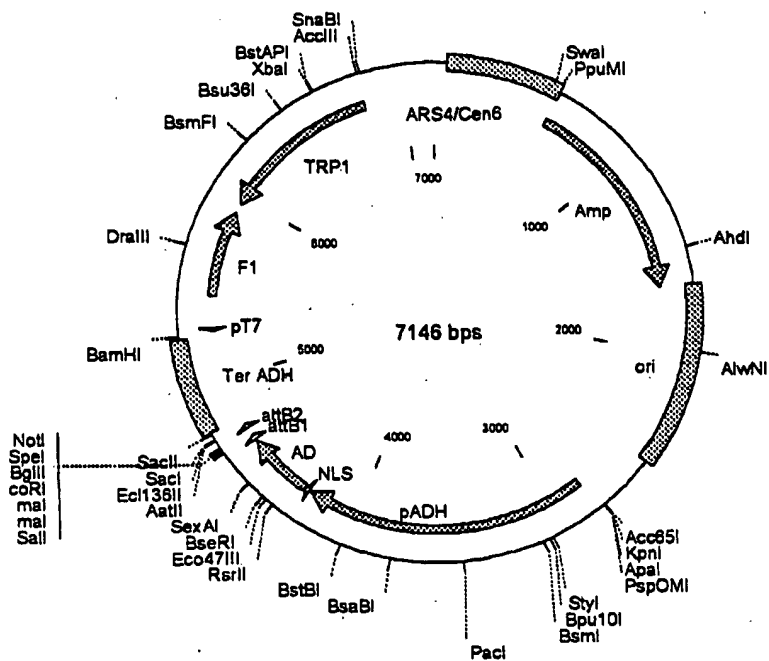


FIGURE 99A

pMAB86 7146 bp

GACGAAAGGGCCTCGTGATACGCCTATTTTTATAGGTTAATGTCATGATAATAATGGTTT
CTTAGGACGGATCGCTTGCCTGTAACCTACACGCGCCTCGTATCTTTAATGATGGAATA
ATTTGGGAATTTACTCTGTGTTTATTTATTTATGTTTTGATTTTGGATTTTAGAAAGT
AAATAAAGAAGGTAGAAGAGTTACGGAATGAAGAAAAAATAAACAAAGGTTTAAAAA
ATTTCAACAAAAAGCGTACTTTACATATATATTTATTAGACAAGAAAAGCAGATTAAATA
GATATACATTGATTAACGATAAGTAAATGTAAATCACAGGATTTTCGTGTGTGGTCT
TCTACACAGACAAGATGAAACAATTTCGGCATTAATACCTGAGAGCAGGAAGAGCAAGATA
AAAGGTAGTATTTGTTGGCGATCCCCCTAGAGTCTTTTACATCTTCGGAAAAACAAAAC
ATTTTTCTTTAATTTCTTTTTTACTTTCTATTTTTAATTTATATATTTATATTAATAA
GTTAAATTATAATTATTTTATAGCACGTGATGAAAAGGACCCAGGTGGCACTTTTCGG
GGAAATGTGCGCGGAACCCCTATTTGTTTATTTTCTAAATACATTCAAATATGTATCCG
CTCATGAGACAATAACCCCTGATAAATGCTTCAATAATATTGAAAAAGGAAGAGTATGAGT
ATTCAACATTTCCGTGTGCGCCCTTATTCCTTTTTTTCGGGCATTTTGCCTTCCTGTTTT
GCTACCCAGAAACGCTGGTGAAAGTAAAGATGCTGAAGATCAGTTGGGTGCACGAGTG
GGTTACATCGAAGTGGATCTCAACAGCGGTAAAGATCCTTGAGAGTTTTCGCCCCGAAGAA
CGTTTTCCAATGATGAGCACTTTTAAAGTTCGTCTATGTGGCGCGGTATTATCCCGTATT
GACGCCGGGCAAGAGCAACTCGGTGCGCGCATACACTATTCTCAGAATGACTTGGTTGAG
TACTCACCAGCTCACAGAAAAGCATCTTACGGATGGCATGACAGTAAGAGAATTATGCAGT
GCTGCCATAACCATGAGTGATAACACTGCGGCCAACTTACTTCTGACAACGATCGGAGGA
CCGAAGGAGCTAACCGCTTTTTTTCACAACATGGGGGATCATGTAACCTCGCCTTGATCGT
TGGAACCCGAGCTGAATGAAGCCATACCAAACGACGAGCGTGACACCACGATGCCTGTA
GCAATGGCAACAACGTTGCGCAAACTATTAAGTGGCGAACTACTTACTCTAGCTTCCCGG
CAACAATTAATAGACTGGATGGAGGCGGATAAAGTTGCAGGACCACTTCTGCGCTCGGCC
CTTCCGGCTGGCTGGTTTATGTGATAAATCTGGAGCCGGTGAGCGTGGGTCTCGCGGT
ATCTCCAGCAGCTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACAGC
GGCAGTCAGGCAACTATGGATGAACGAAATAGACAGATCGCTGAGATAGGTGCCTCACTG
ATTAAGCATTGGTAACCTGTGACACCAAGTTTACTCATATATACTTTAGATTGATTTAAAA
CTTCATTTTTAATTTAAAGGATCTAGGTGAAGATCCTTTTGATAATCTCATGACCAAA
ATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGA
TCTTCTTGAGATCCTTTTTTCTGCGCGTAATCTGCTGCTTGCAACAAAAAAACCACCG
CTACCAGCGGTGGTTTGTGTCGGGATCAAGAGCTACCAACTCTTTTTCCGAAGGTAAC
GGCTTCAGCGAGCGCAGATACCAATACTGCTCTTCTAGTGTAGCCGTAGTTAGGCCAC
CACTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCTGCTAATCCTGTTACCAGTG
GCTGCTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGTTACCG
GATAAGGCGCAGCGGTGCGGCTGAACGGGGGGTTCGTGCACACAGCCAGCTTGGAGCGA
ACGACCTACACCGAACTGAGATACCTACAGCGTGAGCATTGAGAAAGCGCCACGCTTCCC
GAAGGGAGAAAGGCGGACAGGTATCCGGTAAGCGGCAGGGTCGGAACAGGAGAGCGCACG
AGGGAGCTTCCAGGGGGGAACGCCTGGTATCTTTATAGTCCTGTGCGGTTTCGCCACCTC
TGACTTGAGCGTCGATTTTGTGATGCTCGTCAGGGGGGCGAGCCTATGGAAAAACGCC
AGCAACGCGGCCTTTTTTACGGTTCCTGGCCTTTTGCTGGCCTTTTGCTCACATGTTCTTT
CCTGCGTTATCCCCTGATTCTGTGGATAACCGTATTACCGCCTTTGAGTGAGCTGATACC
GCTCGCCGCGAGCCGAACGACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCGGAAGAGCGC
CCAATACGCAAAACCGCTCTCCCCGCGCGTTGGCCGATTCATTAATGCAGCTGGCACGAC
AGGTTTCCCGACTGGAAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTACCTCACT
CATTAGGCACCCCAGGCTTTACACTTTATGCTTCCGGCTCCTATGTTGTGTGGAATTGTG
AGCGGATAACAATTTACACACAGGAAACAGCTATGACCATGATTACGCCAAGCTCGGAATT
AACCCTCACTAAAGGGAACAAAAGCTGGGTACCGGGCCCCCCTCGAGATCCGGGATCGA
AGAAATGATGGTAAATGAAATAGGAAATCAAGGAGCATGAAGGCAAAAGACAAATATAAG
GGTCGAACGAAAAATAAAGTGAAAAGTGTGATATGATGTATTTGGCTTTGCGGCGCCGA
AAAAACGAGTTTACGCAATTGCACAATCATGCTGACTCTGTGGCGGACCCGCGCTCTTGC
CGGCCCGCGGATAACGCTGGGCGTGAGGCTGTGCCCGGCGGAGTTTTTGTGCGCTGCATT
TTCCAAGTTTACCCTGCGCTAAGGGGCGAGATTGAGAAGCAATAAGAATGCCGGTTGG
GGTTCGATGATGACGACCAAGCAAGTGGTGTGCTATTATTTAAGTTGCCGAAAGAACCTG
AGTGCATTTGCAACATGAGTATACTAGAAGAATGAGCCAAGACTTGCGAGACGCGAGTTT
GCCGGTGGTGCGAACAAATAGAGCGACCATGACCTTGAAGGTGAGACGCGCATAACCGCTA-

FIGURE 99B

GAGTACTTTGAAGAGGAAACAGCAATAGGGTTGCTACCAGTATAAATAGACAGGTACATA
CAACACTGGAAATGGTTGTCTGTTTGAGTACGCTTTCAATTCATTTGGGTGTGCACTTTA
TTATGTTACAATATGGAAGGGAACCTTTACACTTCTCCTATGCACATATATTAATTAAAGT
CCAATGCTAGTAGAGAAGGGGGGTAACACCCCTCCGCGCTCTTTTCCGATTTTTTTCTAA
ACCGTGGAAATATTTCCGATATCCTTTTGTGTTTCCGGGTGTACAATATGGACTTCCTCT
TTTCTGGCAACCAACCCATACATCGGGATTCCCTATAATACCTTCGTTGGTCTCCCTAAC
ATGTAGGTGGCGGAGGGGAGATATACAATAGAACAGATACCAGACAAGACATAATGGGCT
AAACAAGACTACACCAATTACACTGCCTCATTGATGGTGGTACATAACGAACATAACTG
TAGCCCTAGACTTGATAGCCATCATCATATCGAAGTTTCACTACCCTTTTTCCATTTGCC
ATCTATTGAAGTAATAATAGGCGCATGCAACTTCTTTTCTTTTTTTTTCTTTCTCTCTC
CCCCGTTGTTGTCTCACCATATCCGCAATGACAAAAAATGATGGAAGACACTAAAGGA
AAAAATTAAACGACAAAGACAGCACCAACAGATGTCGTTGTTCCAGAGCTGATGAGGGGTA
TCTTCGAACACGAACTTTTTCCTTCCTTATTACGCACTACTCTCTAATGAGCA
ACGGTATACGGCCTTCTTCCAGTTACTTGAATTTGAAATAAAAAAGTTTGCCGCTTTG
CTATCAAGTATAAATAGACCTGCAATTATTAATCTTTTGTTCCTCGTCATTGTTCTCGT
TCCCTTTCTTCTTGTCTTTCTTTTCTGCACAATATTTCAAGCTATACCAAGCATACAATC
AACTCCAAGCTTATGCCCAAGAAGAAGCGGAAGGTCTCGAGCGGCGCCAATTTAATCAA
AGTGGGAATATTGCTGATAGCTCATTGTCTTCACTTCACTAACAGTAGCAACGGTCCG
AACCTCATACAACTCAAAACAAATTTCAAGCGCTTTCACAACCAATTGCCTCCTCTAAC
GTTTCATGATACTTATGAATAATGAAATCAGGCTAGTAAAATTGATGATGGTAATAAT
TCAAAAACCACTGTCACTGGTTGGACGGACCAAACTGCGTATAACGCGTTTGAATCACT
ACAGGGATGTTAATAACCACTACAATGGATGATGTATATACTATCTATTGATGATGAA
GATACCCACCAAAACCAAAAAAGAGGGTGGGTGCGATCACAAGTTTGTACAAAAAGCA
GGCTTGTGACCCCGGGAATTCAGATCTACTAGTGCGGCGCGCACGCGTACCCAGCTTTCT
TGTACAAAGTGGTGACGTCGAGCTCTAAGTAAGTAACGGCGCGCCACCGCGGTGGAGCTTT
GGACTTCTTCGCCAGAGGTTTGGTCAAGTCTCCAATCAAGTTTGTGCGCTTGTCTACCTT
GCCAGAAATTTACGAAAAGATGGAAAAGGGTCAAATCGTTGGTAGATACGTTGTTGACAC
TTCTAAATAAGCGAATTTCTTATGATTTATGATTTTTATTATTAAATAAGTTATAAAAA
AATAAGTGTATACAAATTTTAAAGTGAATCTTAGGTTTTAAACGAAAATTTCTGTTCTT
GAGTAATCTTTTCTGTAGGTGAGTTGCTTTCTCAGGTATAGCATGAGGTGCTCTTAT
TGACCACACCTCTACCGGCATGCCGAGCAAATGCCTGCAAATCGCTCCCCATTTACCCA
ATTGTAGATATGCTAACTCCAGCAATGAGTTGATGAATCTCGGTGTGTATTTTATGTCTT
CAGAGGACAATACCTGTTGTAATCGTTCTTCCACACGGATCCCAATTCGCCCTATGAGTA
GTCGTATTACAATTCAGTGGCGTCTGTTTACAACGTCGTGACTGGGAAAACCCCTGGCGT
TACCCAACCTTAATCGCCTTGACGACATCCCCCTTTGCCAGCTGGCGTAATAGCGAAGA
GGCCCGCACCGATCGCCCTTCCCAACAGTTGCGCAGCCTGAATGGCGAATGGACGCGCCC
TGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGGTTACGCGCAGCGTGACCGCTACACTT
GCCAGCGCCCTAGCGCCGCTCCTTTGCTTTCTTCCCTTCTTCTCGCCACGTTGCGC
GGCTTTCCCCGTCAAGCTCTAAATCGGGGGCTCCCTTTAGGGTTCCGATTTAGTGCTTTA
CGGCACCTCGACCCCAAAAACTTGATTAGGGTGATGGTTACGTAAGTGGGCCATCGCCC
TGATAGACGGTTTTTCGCCCTTTGACGTTGGAGTCCACGTTCTTTAATAGTGGACTCTTG
TTCCAAACTGGAACAACACTCAACCTATCTCGGTCTATTCTTTTGATTTATAAGGGATT
TTGCCGATTTGCGCCTATTGGTTAAAAAATGAGCTGATTTAACAAAAATTTAACGCGAAT
TTTAACAAAATATTAACGTTTACAATTTCTGATGCGGTATTTTCTCCTTACGCATCTGT
GCGGTATTTACACCGCAGGCAAGTGCACAAACAATACTTAAATAAATACTACTCAGTAA
TAACCTATTTCTTAGCATTTTGTACGAAATTTGCTATTTTGTAGAGTCTTTTACACCAT
TTGTCTCCACACCTCCGCTTACATCAACACCAATAACGCCATTTAATCTAAGCGCATCAC
CAACATTTTCTGGCGTCAGTCCACCAGCTAACATAAAATGTAAGCTTTCGGGGCTCTCTT
GCCTTCCAACCCAGTCAGAAATCGAGTTCCAATCCAAAAGTTTACCTGTCCACCTGCTT
CTGAATCAAAACAGGGAATAAACGAATGAGGTTTCTGTGAAGCTGCACTGAGTAGTATGT
TGCAGTCTTTTGGAAATACGAGTCTTTTAATAACTGGCAAACCGAGGAACCTTTGGTATT
CTTGCCACGACTCATCTCCATGCAGTTGGACGATATCAATGCCGTAATCATTGACCAGAG
CCAAAACATCCTCCTTAGGTTGATTACGAAACACGCCAACCAAGTATTTCCGAGTGCCCTG
AACTATTTTATATGCTTTTACAAGACTTGAAATTTTCTTGCAATAACCGGGTCAATTG
TTCTCTTCTATTTGGGCACACATATAATACCCAGCAAGTCAGCATCGGAATCTAGAGCAC
ATTTCTGCGCCTCTGTGCTCTGCAAGCCGCAACTTTACCAATGGACCAGAACTACCTG
TGAAATTAATAACAGACATACTCCAAGCTGCCTTTGTGTGCTTAATCACGTATACTCACG
TGCTCAATAGTCACCAATGCCCTCCCTCTTGGCCCTCTCCTTTTCTTTTTTCGACCGAAT-

TAATTCTTAATCGGCAAAAAAGAAAAGCTCCGGATCAAGATTGTACGTAAGGTGACAAG
CTATTTTTCAATAAAGAATATCTTCCACTACTGCCATCTGGCGTCATAACTGCAAAGTAC
ACATATATTACGATGCTGTCTATTAAATGCTTCCTATATTATATATATAGTAATGTCGTT
TATGGTGCACTCTCAGTACAATCTGCTCTGATGCCGCATAGTTAAGCCAGCCCCGACACC
CGCCAACACCCGCTGACGCGCCCTGACGGGCTTGTCTGCTCCCGGCATCCGCTTACAGAC
AAGCTGTGACCGTCTCCGGGAGCTGCATGTGTCAGAGGTTTTACCGTCATCACCGAAAC
GCGCGA

FIGURE 99D

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM
(PCT Rule 13bis)

REC'D

A. The indications made below relate to the microorganism referred to in the description on page <u>54</u> , line <u>8</u> .	
B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input checked="" type="checkbox"/>	
Name of depositary institution Agricultural Research Culture Collection (NRRL) International Depository Authority	
Address of depositary institution (including postal code and country) 1815 N. University Street Peoria, Illinois 61604 United States of America	
Date of deposit February 27, 1999	Accession Number NRRL B-30103
C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is continued on an additional sheet <input type="checkbox"/>	
Escherichia coli DB3.1(pEZC15101) In respect of those designations in which a European Patent is sought a sample of the deposited microorganism will be made available until the publication of the mention of the grant of the European patent or until the date on which the application has been refused or withdrawn or is deemed to be withdrawn, only by the issue of such a sample to an expert nominated by the person requesting the sample (Rule 28(4) EPC).	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)	
E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)	
The indications listed below will be submitted to the international Bureau later (specify the general nature of the indications, e.g., "Accession Number of Deposit")	

<p style="text-align: center;">For receiving Office use only</p> <p><input checked="" type="checkbox"/> This sheet was received with the international application</p> <p>Authorized officer <i>B. Indici</i></p>	<p style="text-align: center;">For International Bureau use only</p> <p><input type="checkbox"/> This sheet was received by the International Bureau on:</p> <p>Authorized officer</p>
---	--

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM
(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page <u>55</u> , line <u>16</u> .	
B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input checked="" type="checkbox"/>	
Name of depositary institution Agricultural Research Culture Collection (NRRL) International Depository Authority	
Address of depositary institution (including postal code and country) 1815 N. University Street Peoria, Illinois 61604 United States of America	
Date of deposit February 27, 1999	Accession Number NRRL B-30100
C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is continued on an additional sheet <input type="checkbox"/>	
Escherichia coli DB3.1(pENTR-1A) In respect of those designations in which a European Patent is sought a sample of the deposited microorganism will be made available until the publication of the mention of the grant of the European patent or until the date on which the application has been refused or withdrawn or is deemed to be withdrawn, only by the issue of such a sample to an expert nominated by the person requesting the sample (Rule 28(4) EPC).	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (If the indications are not for all designated States)	
E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)	
The indications listed below will be submitted to the international Bureau later (specify the general nature of the indications, e.g., "Accession Number of Deposit")	

<p align="center">For receiving Office use only</p> <p><input checked="" type="checkbox"/> This sheet was received with the international application</p> <p>Authorized officer <i>B. Fudis</i></p>	<p align="center">For International Bureau use only</p> <p><input type="checkbox"/> This sheet was received by the International Bureau on:</p> <p>Authorized officer</p>
--	--


INDICATIONS RELATING TO A DEPOSITED MICROORGANISM
(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page <u>55</u> , line <u>16</u> .	
B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input checked="" type="checkbox"/>	
Name of depositary institution Agricultural Research Culture Collection (NRRL) International Depository Authority	
Address of depositary institution (including postal code and country) 1815 N. University Street Peoria, Illinois 61604 United States of America	
Date of deposit February 27, 1999	Accession Number NRRL B-30102
C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is continued on an additional sheet <input type="checkbox"/>	
Escherichia coli DB3.1(pENTR-3C) In respect of those designations in which a European Patent is sought a sample of the deposited microorganism will be made available until the publication of the mention of the grant of the European patent or until the date on which the application has been refused or withdrawn or is deemed to be withdrawn, only by the issue of such a sample to an expert nominated by the person requesting the sample (Rule 28(4) EPC).	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)	
E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)	
The indications listed below will be submitted to the international Bureau later (specify the general nature of the indications, e.g., "Accession Number of Deposit")	

For receiving Office use only	For International Bureau use only
<input checked="" type="checkbox"/> This sheet was received with the international application	<input type="checkbox"/> This sheet was received by the International Bureau on:
Authorized officer <i>B. Indie</i>	Authorized officer

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM
(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page <u>55</u> , line <u>16</u> .	
B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input checked="" type="checkbox"/>	
Name of depositary institution Agricultural Research Culture Collection (NRRL) International Depository Authority	
Address of depositary institution (including postal code and country) 1815 N. University Street Peoria, Illinois 61604 United States of America	
Date of deposit February 27, 1999	Accession Number NRRL B-30101
C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is continued on an additional sheet <input type="checkbox"/>	
Escherichia coli DB3.1(pENTR-2B) In respect of those designations in which a European Patent is sought a sample of the deposited microorganism will be made available until the publication of the mention of the grant of the European patent or until the date on which the application has been refused or withdrawn or is deemed to be withdrawn, only by the issue of such a sample to an expert nominated by the person requesting the sample (Rule 28(4) EPC).	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)	
E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)	
The indications listed below will be submitted to the international Bureau later (specify the general nature of the indications, e.g., "Accession Number of Deposit")	

For receiving Office use only	For International Bureau use only
<input checked="" type="checkbox"/> This sheet was received with the international application	<input type="checkbox"/> This sheet was received by the International Bureau on:
Authorized officer 	Authorized officer

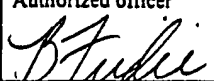
INDICATIONS RELATING TO A DEPOSITED MICROORGANISM
(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page <u>WIPO</u> <u>PCT</u> <u>20-21</u>	
B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input checked="" type="checkbox"/>	
Name of depositary institution Agricultural Research Culture Collection (NRRL) International Depositary Authority	
Address of depositary institution (including postal code and country) 1815 N. University Street Peoria, Illinois 61604 United States of America	
Date of deposit February 27, 1999	Accession Number NRRL B-30108
C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is continued on an additional sheet <input type="checkbox"/>	
Escherichia coli DB10B(pCMVSPORT6) In respect of those designations in which a European Patent is sought a sample of the deposited microorganism will be made available until the publication of the mention of the grant of the European patent or until the date on which the application has been refused or withdrawn or is deemed to be withdrawn, only by the issue of such a sample to an expert nominated by the person requesting the sample (Rule 28(4) EPC).	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)	
E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)	
The indications listed below will be submitted to the international Bureau later (specify the general nature of the indications, e.g., "Accession Number of Deposit")	

<p align="center">For receiving Office use only</p> <div style="border: 1px solid black; padding: 5px; margin-bottom: 5px;"> <input checked="" type="checkbox"/> This sheet was received with the international application </div> <div style="border: 1px solid black; padding: 5px;"> Authorized officer </div>	<p align="center">For International Bureau use only</p> <div style="border: 1px solid black; padding: 5px; margin-bottom: 5px;"> <input type="checkbox"/> This sheet was received by the International Bureau on: </div> <div style="border: 1px solid black; padding: 5px;"> Authorized officer </div>
---	--


INDICATIONS RELATING TO A DEPOSITED MICROORGANISM
(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page <u>54</u> , line <u>9</u> .	
B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input checked="" type="checkbox"/>	
Name of depositary institution Agricultural Research Culture Collection (NRRL) International Depository Authority	
Address of depositary institution (including postal code and country) 1815 N. University Street Peoria, Illinois 61604 United States of America	
Date of deposit February 27, 1999	Accession Number NRRL B-30105
C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is continued on an additional sheet <input type="checkbox"/>	
Escherichia coli DB3.1(pEZC15103) In respect of those designations in which a European Patent is sought a sample of the deposited microorganism will be made available until the publication of the mention of the grant of the European patent or until the date on which the application has been refused or withdrawn or is deemed to be withdrawn, only by the issue of such a sample to an expert nominated by the person requesting the sample (Rule 28(4) EPC).	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (If the indications are not for all designated States)	
E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)	
The indications listed below will be submitted to the international Bureau later (specify the general nature of the indications, e.g., "Accession Number of Deposit")	

For receiving Office use only	For International Bureau use only
<input checked="" type="checkbox"/> This sheet was received with the international application	<input type="checkbox"/> This sheet was received by the International Bureau on:
Authorized officer 	Authorized officer

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM
(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page <u>54</u> , line <u>9</u> .	
B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input checked="" type="checkbox"/>	
Name of depositary institution Agricultural Research Culture Collection (NRRL) International Depositary Authority	
Address of depositary institution (including postal code and country) 1815 N. University Street Peoria, Illinois 61604 United States of America	
Date of deposit February 27, 1999	Accession Number NRRL B-30104
C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is continued on an additional sheet <input type="checkbox"/>	
Escherichia coli DB3.1(pEZC15102) In respect of those designations in which a European Patent is sought a sample of the deposited microorganism will be made available until the publication of the mention of the grant of the European patent or until the date on which the application has been refused or withdrawn or is deemed to be withdrawn, only by the issue of such a sample to an expert nominated by the person requesting the sample (Rule 28(4) EPC).	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)	
E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)	
The indications listed below will be submitted to the international Bureau later (specify the general nature of the indications, e.g., "Accession Number of Deposit")	

For receiving Office use only	For International Bureau use only
<input checked="" type="checkbox"/> This sheet was received with the international application	<input type="checkbox"/> This sheet was received by the International Bureau on:
Authorized officer 	Authorized officer

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM
(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page <u>52</u> , line <u>31</u> .	
B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input checked="" type="checkbox"/>	
Name of depositary institution Agricultural Research Culture Collection (NRRL) International Depository Authority	
Address of depositary institution (including postal code and country) 1815 N. University Street Peoria, Illinois 61604 United States of America	
Date of deposit February 27, 1999	Accession Number NRRL B-30099
C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is continued on an additional sheet <input type="checkbox"/>	
Escherichia coli DB3.1(pAHPKan) or Escherichia coli DB3.1(pAttPKan) In respect of those designations in which a European Patent is sought a sample of the deposited microorganism will be made available until the publication of the mention of the grant of the European patent or until the date on which the application has been refused or withdrawn or is deemed to be withdrawn, only by the issue of such a sample to an expert nominated by the person requesting the sample (Rule 28(4) EPC).	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (If the indications are not for all designated States)	
E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)	
The indications listed below will be submitted to the international Bureau later (specify the general nature of the indications, e.g., "Accession Number of Deposit")	

<p align="center">For receiving Office use only</p> <p><input checked="" type="checkbox"/> This sheet was received with the international application</p> <p>Authorized officer Barbara Fridle PCT Operations - IPD Team 1 703) 305-3230 (FA)</p>	<p align="center">For International Bureau use only</p> <p><input type="checkbox"/> This sheet was received by the International Bureau on:</p> <p>Authorized officer</p>
--	---

Escherichia coli DB3.1(pENTR-3C)**ICELAND**

The applicant hereby requests that, until the application has been laid open to public inspection (by the Icelandic Patent Office), or has been finally decided upon by the Icelandic Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art.

NETHERLANDS

The applicant hereby requests that until the date of a grant of a Netherlands patent or until the date on which the application is refused or withdrawn or lapsed, the microorganism shall be made available as provided in Rule 31F(1) of the Patent Rules only by the issue of a sample to an expert. The request to this effect must be furnished by the applicant with the Netherlands Industrial Property Office before the date on which the application is made available to the public under Section 22C or Section 25 of the Patents Act of the Kingdom of the Netherlands, whichever of the two dates occurs earlier.

NORWAY

The applicant hereby requests that, until the application has been laid open to public inspection (by the Norwegian Patent Office), or has been finally decided upon by the Norwegian Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Norwegian Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Norwegian Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Norwegian Patent office or any person approved by the applicant in the individual case.

SINGAPORE

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for international publication of the application.

Escherichia coli DB3.1(pENTR-3C)**SWEDEN**

The applicant hereby requests that, until the application has been laid open to public inspection (by the Swedish Patent Office), or has been finally decided upon by the Swedish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Swedish Patent office or any person approved by the applicant in the individual case.

UNITED KINGDOM

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for international publication of the application.

Escherichia coli DB3.1(pENTR-2B)**AUSTRALIA**

The applicant hereby gives notice that the furnishing of a sample of a microorganism shall only be effected prior to the grant of a patent, or prior to the lapsing, refusal or withdrawal of the application, to a person who is a skilled addressee without an interest in the invention (Regulation 3.25(3) of the Australian Patents Regulations).

CANADA

The applicant hereby requests that, until either a Canadian patent has been issued on the basis of the application or the application has been refused, or is abandoned and no longer subject to reinstatement, or is withdrawn, the furnishing of a sample of deposited biological material referred to in the application only be effected to an independent expert nominated by the Commissioner of Patents.

DENMARK

The applicant hereby requests that, until the application has been laid open to public inspection (by the Danish Patent Office), or has been finally decided upon by the Danish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Danish Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Danish Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Danish Patent office or any person approved by the applicant in the individual case.

FINLAND

The applicant hereby requests that, until the application has been laid open to public inspection (by the National Board of Patents and Registration), or has been finally decided upon by the National Board of Patents and Registration without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the National Board of Patents and Registration or any person approved by the applicant in the individual case.

Escherichia coli DB3.1(pENTR-2B)**ICELAND**

The applicant hereby requests that, until the application has been laid open to public inspection (by the Icelandic Patent Office), or has been finally decided upon by the Icelandic Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art.

NETHERLANDS

The applicant hereby requests that until the date of a grant of a Netherlands patent or until the date on which the application is refused or withdrawn or lapsed, the microorganism shall be made available as provided in Rule 31F(1) of the Patent Rules only by the issue of a sample to an expert. The request to this effect must be furnished by the applicant with the Netherlands Industrial Property Office before the date on which the application is made available to the public under Section 22C or Section 25 of the Patents Act of the Kingdom of the Netherlands, whichever of the two dates occurs earlier.

NORWAY

The applicant hereby requests that, until the application has been laid open to public inspection (by the Norwegian Patent Office), or has been finally decided upon by the Norwegian Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Norwegian Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Norwegian Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Norwegian Patent office or any person approved by the applicant in the individual case.

SINGAPORE

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for international publication of the application.

Escherichia coli DB3.1(pENTR-2B)**SWEDEN**

The applicant hereby requests that, until the application has been laid open to public inspection (by the Swedish Patent Office), or has been finally decided upon by the Swedish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Swedish Patent office or any person approved by the applicant in the individual case.

UNITED KINGDOM

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for international publication of the application.

Escherichia coli DB3.1(pENTR-1A)**AUSTRALIA**

The applicant hereby gives notice that the furnishing of a sample of a microorganism shall only be effected prior to the grant of a patent, or prior to the lapsing, refusal or withdrawal of the application, to a person who is a skilled addressee without an interest in the invention (Regulation 3.25(3) of the Australian Patents Regulations).

CANADA

The applicant hereby requests that, until either a Canadian patent has been issued on the basis of the application or the application has been refused, or is abandoned and no longer subject to reinstatement, or is withdrawn, the furnishing of a sample of deposited biological material referred to in the application only be effected to an independent expert nominated by the Commissioner of Patents.

DENMARK

The applicant hereby requests that, until the application has been laid open to public inspection (by the Danish Patent Office), or has been finally decided upon by the Danish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Danish Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Danish Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Danish Patent office or any person approved by the applicant in the individual case.

FINLAND

The applicant hereby requests that, until the application has been laid open to public inspection (by the National Board of Patents and Registration), or has been finally decided upon by the National Board of Patents and Registration without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the National Board of Patents and Registration or any person approved by the applicant in the individual case.

Escherichia coli DB3.1(pENTR-1A)**ICELAND**

The applicant hereby requests that, until the application has been laid open to public inspection (by the Icelandic Patent Office), or has been finally decided upon by the Icelandic Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art.

NETHERLANDS

The applicant hereby requests that until the date of a grant of a Netherlands patent or until the date on which the application is refused or withdrawn or lapsed, the microorganism shall be made available as provided in Rule 31F(1) of the Patent Rules only by the issue of a sample to an expert. The request to this effect must be furnished by the applicant with the Netherlands Industrial Property Office before the date on which the application is made available to the public under Section 22C or Section 25 of the Patents Act of the Kingdom of the Netherlands, whichever of the two dates occurs earlier.

NORWAY

The applicant hereby requests that, until the application has been laid open to public inspection (by the Norwegian Patent Office), or has been finally decided upon by the Norwegian Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Norwegian Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Norwegian Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Norwegian Patent office or any person approved by the applicant in the individual case.

SINGAPORE

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for international publication of the application.

Escherichia coli DB3.1(pENTR-1A)**SWEDEN**

The applicant hereby requests that, until the application has been laid open to public inspection (by the Swedish Patent Office), or has been finally decided upon by the Swedish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Swedish Patent office or any person approved by the applicant in the individual case.

UNITED KINGDOM

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for international publication of the application.

*Escherichia coli DB10B(pCMVSPORT6)***AUSTRALIA**

The applicant hereby gives notice that the furnishing of a sample of a microorganism shall only be effected prior to the grant of a patent, or prior to the lapsing, refusal or withdrawal of the application, to a person who is a skilled addressee without an interest in the invention (Regulation 3.25(3) of the Australian Patents Regulations).

CANADA

The applicant hereby requests that, until either a Canadian patent has been issued on the basis of the application or the application has been refused, or is abandoned and no longer subject to reinstatement, or is withdrawn, the furnishing of a sample of deposited biological material referred to in the application only be effected to an independent expert nominated by the Commissioner of Patents.

DENMARK

The applicant hereby requests that, until the application has been laid open to public inspection (by the Danish Patent Office), or has been finally decided upon by the Danish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Danish Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Danish Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Danish Patent office or any person approved by the applicant in the individual case.

FINLAND

The applicant hereby requests that, until the application has been laid open to public inspection (by the National Board of Patents and Registration), or has been finally decided upon by the National Board of Patents and Registration without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the National Board of Patents and Registration or any person approved by the applicant in the individual case.

Escherichia coli DB3.1(pAHPKan) or *Escherichia coli* DB3.1(pAttPKan)

AUSTRALIA

The applicant hereby gives notice that the furnishing of a sample of a microorganism shall only be effected prior to the grant of a patent, or prior to the lapsing, refusal or withdrawal of the application, to a person who is a skilled addressee without an interest in the invention (Regulation 3.25(3) of the Australian Patents Regulations).

CANADA

The applicant hereby requests that, until either a Canadian patent has been issued on the basis of the application or the application has been refused, or is abandoned and no longer subject to reinstatement, or is withdrawn, the furnishing of a sample of deposited biological material referred to in the application only be effected to an independent expert nominated by the Commissioner of Patents.

DENMARK

The applicant hereby requests that, until the application has been laid open to public inspection (by the Danish Patent Office), or has been finally decided upon by the Danish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Danish Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Danish Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Danish Patent office or any person approved by the applicant in the individual case.

FINLAND

The applicant hereby requests that, until the application has been laid open to public inspection (by the National Board of Patents and Registration), or has been finally decided upon by the National Board of Patents and Registration without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the National Board of Patents and Registration or any person approved by the applicant in the individual case.

Escherichia coli DB3.1(pAHPKan) or Escherichia coli DB3.1(pAttPKan)

ICELAND

The applicant hereby requests that, until the application has been laid open to public inspection (by the Icelandic Patent Office), or has been finally decided upon by the Icelandic Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art.

NETHERLANDS

The applicant hereby requests that until the date of a grant of a Netherlands patent or until the date on which the application is refused or withdrawn or lapsed, the microorganism shall be made available as provided in Rule 31F(1) of the Patent Rules only by the issue of a sample to an expert. The request to this effect must be furnished by the applicant with the Netherlands Industrial Property Office before the date on which the application is made available to the public under Section 22C or Section 25 of the Patents Act of the Kingdom of the Netherlands, whichever of the two dates occurs earlier.

NORWAY

The applicant hereby requests that, until the application has been laid open to public inspection (by the Norwegian Patent Office), or has been finally decided upon by the Norwegian Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Norwegian Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Norwegian Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Norwegian Patent office or any person approved by the applicant in the individual case.

SINGAPORE

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for international publication of the application.

Escherichia coli DB3.1(pAHPKan) or Escherichia coli DB3.1(pAttPKa ϕ)

SWEDEN

The applicant hereby requests that, until the application has been laid open to public inspection (by the Swedish Patent Office), or has been finally decided upon by the Swedish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Swedish Patent office or any person approved by the applicant in the individual case.

UNITED KINGDOM

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for international publication of the application.

*Escherichia coli DB10B(pCMVSpot6)***ICELAND**

The applicant hereby requests that, until the application has been laid open to public inspection (by the Icelandic Patent Office), or has been finally decided upon by the Icelandic Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art.

NETHERLANDS

The applicant hereby requests that until the date of a grant of a Netherlands patent or until the date on which the application is refused or withdrawn or lapsed, the microorganism shall be made available as provided in Rule 31F(1) of the Patent Rules only by the issue of a sample to an expert. The request to this effect must be furnished by the applicant with the Netherlands Industrial Property Office before the date on which the application is made available to the public under Section 22C or Section 25 of the Patents Act of the Kingdom of the Netherlands, whichever of the two dates occurs earlier.

NORWAY

The applicant hereby requests that, until the application has been laid open to public inspection (by the Norwegian Patent Office), or has been finally decided upon by the Norwegian Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Norwegian Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Norwegian Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Norwegian Patent office or any person approved by the applicant in the individual case.

SINGAPORE

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for international publication of the application.

Escherichia coli DB10B(pCMVSPORT6)

SWEDEN

The applicant hereby requests that, until the application has been laid open to public inspection (by the Swedish Patent Office), or has been finally decided upon by the Swedish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Swedish Patent office or any person approved by the applicant in the individual case.

UNITED KINGDOM

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for international publication of the application.

Escherichia coli DB3.1(pEZC15103)**AUSTRALIA**

The applicant hereby gives notice that the furnishing of a sample of a microorganism shall only be effected prior to the grant of a patent, or prior to the lapsing, refusal or withdrawal of the application, to a person who is a skilled addressee without an interest in the invention (Regulation 3.25(3) of the Australian Patents Regulations).

CANADA

The applicant hereby requests that, until either a Canadian patent has been issued on the basis of the application or the application has been refused, or is abandoned and no longer subject to reinstatement, or is withdrawn, the furnishing of a sample of deposited biological material referred to in the application only be effected to an independent expert nominated by the Commissioner of Patents.

DENMARK

The applicant hereby requests that, until the application has been laid open to public inspection (by the Danish Patent Office), or has been finally decided upon by the Danish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Danish Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Danish Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Danish Patent office or any person approved by the applicant in the individual case.

FINLAND

The applicant hereby requests that, until the application has been laid open to public inspection (by the National Board of Patents and Registration), or has been finally decided upon by the National Board of Patents and Registration without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the National Board of Patents and Registration or any person approved by the applicant in the individual case.

Escherichia coli DB3.1(pEZC15103)

ICELAND

The applicant hereby requests that, until the application has been laid open to public inspection (by the Icelandic Patent Office), or has been finally decided upon by the Icelandic Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art.

NETHERLANDS

The applicant hereby requests that until the date of a grant of a Netherlands patent or until the date on which the application is refused or withdrawn or lapsed, the microorganism shall be made available as provided in Rule 31F(1) of the Patent Rules only by the issue of a sample to an expert. The request to this effect must be furnished by the applicant with the Netherlands Industrial Property Office before the date on which the application is made available to the public under Section 22C or Section 25 of the Patents Act of the Kingdom of the Netherlands, whichever of the two dates occurs earlier.

NORWAY

The applicant hereby requests that, until the application has been laid open to public inspection (by the Norwegian Patent Office), or has been finally decided upon by the Norwegian Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Norwegian Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Norwegian Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Norwegian Patent office or any person approved by the applicant in the individual case.

SINGAPORE

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for international publication of the application.

Escherichia coli DB3.1(pEZC15103)**SWEDEN**

The applicant hereby requests that, until the application has been laid open to public inspection (by the Swedish Patent Office), or has been finally decided upon by the Swedish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Swedish Patent office or any person approved by the applicant in the individual case.

UNITED KINGDOM

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for international publication of the application.

Escherichia coli DB3.1(pEZC15102)**AUSTRALIA**

The applicant hereby gives notice that the furnishing of a sample of a microorganism shall only be effected prior to the grant of a patent, or prior to the lapsing, refusal or withdrawal of the application, to a person who is a skilled addressee without an interest in the invention (Regulation 3.25(3) of the Australian Patents Regulations).

CANADA

The applicant hereby requests that, until either a Canadian patent has been issued on the basis of the application or the application has been refused, or is abandoned and no longer subject to reinstatement, or is withdrawn, the furnishing of a sample of deposited biological material referred to in the application only be effected to an independent expert nominated by the Commissioner of Patents.

DENMARK

The applicant hereby requests that, until the application has been laid open to public inspection (by the Danish Patent Office), or has been finally decided upon by the Danish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Danish Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Danish Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Danish Patent office or any person approved by the applicant in the individual case.

FINLAND

The applicant hereby requests that, until the application has been laid open to public inspection (by the National Board of Patents and Registration), or has been finally decided upon by the National Board of Patents and Registration without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the National Board of Patents and Registration or any person approved by the applicant in the individual case.

Escherichia coli DB3.1(pEZC15102)

ICELAND

The applicant hereby requests that, until the application has been laid open to public inspection (by the Icelandic Patent Office), or has been finally decided upon by the Icelandic Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art.

NETHERLANDS

The applicant hereby requests that until the date of a grant of a Netherlands patent or until the date on which the application is refused or withdrawn or lapsed, the microorganism shall be made available as provided in Rule 31F(1) of the Patent Rules only by the issue of a sample to an expert. The request to this effect must be furnished by the applicant with the Netherlands Industrial Property Office before the date on which the application is made available to the public under Section 22C or Section 25 of the Patents Act of the Kingdom of the Netherlands, whichever of the two dates occurs earlier.

NORWAY

The applicant hereby requests that, until the application has been laid open to public inspection (by the Norwegian Patent Office), or has been finally decided upon by the Norwegian Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Norwegian Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Norwegian Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Norwegian Patent office or any person approved by the applicant in the individual case.

SINGAPORE

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for international publication of the application.

Escherichia coli DB3.1(pEZC15102)**SWEDEN**

The applicant hereby requests that, until the application has been laid open to public inspection (by the Swedish Patent Office), or has been finally decided upon by the Swedish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Swedish Patent office or any person approved by the applicant in the individual case.

UNITED KINGDOM

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for international publication of the application.

Escherichia coli DB3.1(pEZC15101)**AUSTRALIA**

The applicant hereby gives notice that the furnishing of a sample of a microorganism shall only be effected prior to the grant of a patent, or prior to the lapsing, refusal or withdrawal of the application, to a person who is a skilled addressee without an interest in the invention (Regulation 3.25(3) of the Australian Patents Regulations).

CANADA

The applicant hereby requests that, until either a Canadian patent has been issued on the basis of the application or the application has been refused, or is abandoned and no longer subject to reinstatement, or is withdrawn, the furnishing of a sample of deposited biological material referred to in the application only be effected to an independent expert nominated by the Commissioner of Patents.

DENMARK

The applicant hereby requests that, until the application has been laid open to public inspection (by the Danish Patent Office), or has been finally decided upon by the Danish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Danish Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Danish Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Danish Patent office or any person approved by the applicant in the individual case.

FINLAND

The applicant hereby requests that, until the application has been laid open to public inspection (by the National Board of Patents and Registration), or has been finally decided upon by the National Board of Patents and Registration without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the National Board of Patents and Registration or any person approved by the applicant in the individual case.

Escherichia coli DB3.1(pEZC15101)**ICELAND**

The applicant hereby requests that, until the application has been laid open to public inspection (by the Icelandic Patent Office), or has been finally decided upon by the Icelandic Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art.

NETHERLANDS

The applicant hereby requests that until the date of a grant of a Netherlands patent or until the date on which the application is refused or withdrawn or lapsed, the microorganism shall be made available as provided in Rule 31F(1) of the Patent Rules only by the issue of a sample to an expert. The request to this effect must be furnished by the applicant with the Netherlands Industrial Property Office before the date on which the application is made available to the public under Section 22C or Section 25 of the Patents Act of the Kingdom of the Netherlands, whichever of the two dates occurs earlier.

NORWAY

The applicant hereby requests that, until the application has been laid open to public inspection (by the Norwegian Patent Office), or has been finally decided upon by the Norwegian Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Norwegian Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Norwegian Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Norwegian Patent office or any person approved by the applicant in the individual case.

SINGAPORE

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for international publication of the application.

Escherichia coli DB3.1(pEZC15101)

SWEDEN

The applicant hereby requests that, until the application has been laid open to public inspection (by the Swedish Patent Office), or has been finally decided upon by the Swedish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Swedish Patent office or any person approved by the applicant in the individual case.

UNITED KINGDOM

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for international publication of the application.

Escherichia coli* DB3.1(pENTR-3C)*AUSTRALIA**

The applicant hereby gives notice that the furnishing of a sample of a microorganism shall only be effected prior to the grant of a patent, or prior to the lapsing, refusal or withdrawal of the application, to a person who is a skilled addressee without an interest in the invention (Regulation 3.25(3) of the Australian Patents Regulations).

CANADA

The applicant hereby requests that, until either a Canadian patent has been issued on the basis of the application or the application has been refused, or is abandoned and no longer subject to reinstatement, or is withdrawn, the furnishing of a sample of deposited biological material referred to in the application only be effected to an independent expert nominated by the Commissioner of Patents.

DENMARK

The applicant hereby requests that, until the application has been laid open to public inspection (by the Danish Patent Office), or has been finally decided upon by the Danish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Danish Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Danish Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Danish Patent office or any person approved by the applicant in the individual case.

FINLAND

The applicant hereby requests that, until the application has been laid open to public inspection (by the National Board of Patents and Registration), or has been finally decided upon by the National Board of Patents and Registration without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the National Board of Patents and Registration or any person approved by the applicant in the individual case.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US00/05432

A. CLASSIFICATION OF SUBJECT MATTER

IPC(7) : Please See Extra Sheet.

US CL : 435/91.2, 252.3, 320.1; 530/350; 536/ 23.1, 24.1

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 435/91.2, 252.3, 320.1; 530/350; 536/ 23.1, 24.1

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

NONE

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

Please See Extra Sheet.

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X,P ---- Y,P	US 5,888,732 A (HARTLEY et al.) 30 March 1999, see entire document.	1-21, 25-30 36-38 ----- 22-24, 31-35
X - Y	HASAN et al. Escherichia coli genome targeting, I. Cre-lox-mediated in vitro generation of ori- plasmids and their in vivo chromosomal integration and retrieval. Gene. 1994, Vol. 150, pages 51-56, see entire document.	1-5, 10, 11, 19-21 ----- 15-18, 22-38
X - Y	KATZ et al. Site-specific recombination in Escherichia coli between the att sites of plasmid pSE211 from Saccharopolyspora erythraea. Mol. Gen. Genet. 1991, Vol. 227, pages 155-159, see entire document.	1-11, 19-21 ----- 15-18, 22-38



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents:	*T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
A document defining the general state of the art which is not considered to be of particular relevance	*X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
E earlier document published on or after the international filing date	*Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
L document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	*A* document member of the same patent family
O document referring to an oral disclosure, use, exhibition or other means	
P document published prior to the international filing date but later than the priority date claimed	

Date of the actual completion of the international search

08 MAY 2000

Date of mailing of the international search report

23 MAY 2000

Name and mailing address of the ISA/US
Commissioner of Patents and Trademarks
Box PCT
Washington, D.C. 20231

Facsimile No. (703) 305-3230

Authorized officer

IREM YUCEL

Telephone No. (703) 308-0196

INTERNATIONAL SEARCH REPORTInternational application No.
PCT/US00/05432

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X - Y	ASTUMIAN et al. Site-specific recombination between cloned attP and attB sites from the Haemophilus influenzae bacteriophage HP1 propagated in recombination deficient Escherichia coli. J of Bacteriology. March 1989, Vol. 171, No. 3, pages 1747-1750, see entire document.	1-11, 19-21 ----- 15-18, 22-38

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US00/05432

A. CLASSIFICATION OF SUBJECT MATTER:

IPC (7):

C07H 21/04; C07K 1/00, 14/00; C12N 1/21, 15/00, 15/09, 15/63, 15/70; C12P 19/34

B. FIELDS SEARCHED

Electronic data bases consulted (Name of data base and where practicable terms used):

WEST, STN (CAPLUS); DIALOG (MEDLINE, BIOSIS, SCISEARCH, PASCAL)

Terms: att (B?, P?, R?, L?), MCS, POLYLINKER, PLASMID, VECTOR, LOCALIZATION, SIGNAL, TRANSCRIPTION, TERMIN?, TRANSLATION?, ORI, REPLICON, GST, HEXHIST?, THIOREDOX?, CLEAVAGE, SITE?, SPECIF?, DIRECT?, RECOMBIN?, CLON?, INSERT?

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ BLACK BORDERS
- ☐ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☐ FADED TEXT OR DRAWING
- ☒ BLURRED OR ILLEGIBLE TEXT OR DRAWING
- ☐ SKEWED/SLANTED IMAGES
- ☐ COLOR OR BLACK AND WHITE PHOTOGRAPHS
- ☐ GRAY SCALE DOCUMENTS
- ☐ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☐ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☐ OTHER: _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.